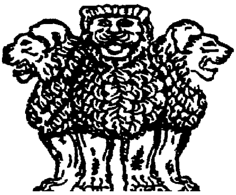

सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI.

I. A. R. I. 6.
MGIPC—S1—6 AR/54—7-7-54—10,000.

OXFORD
ECONOMIC PAPERS

NEW SERIES
VOLUME I
1949

OXFORD
AT THE CLARENDON PRESS
1949

Oxford University Press, Amen House, London E.C. 4
GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON
BOMBAY CALCUTTA MADRAS CAPE TOWN
Geoffrey Cumberlege, Publisher to the University

PRINTED IN GREAT BRITAIN
AT THE UNIVERSITY PRESS, OXFORD
BY CHARLES BATEY, PRINTER TO THE UNIVERSITY

CONTENTS

VOLUME I

Number 1, January 1949

PAGE

THE FUNCTION OF EXCHANGE RATES. *By* SIR HUBERT HENDERSON 1

NOTES ON BRITAIN'S BARGAINING POWER. *By* G. D. A. MACDOUGALL 18

PROGRAMMES AND ALLOCATIONS IN THE PLANNED ECONOMY. *By* T. WILSON 40

A RECONSIDERATION OF THE THEORY OF THE INDIVIDUAL BUSINESS. *By* P. W. S. ANDREWS 54

A REFORMULATION OF THE THEORY OF CONSUMER'S BEHAVIOUR. *By* I. M. D. LITTLE 90

THE NATURE OF INTEREST-RATES. *By* G. L. S. SHACKLE . . 100

TYPES OF COMPETITION AND THE THEORY OF EMPLOYMENT. *By* B. R. WILLIAMS 121

Number 2, June 1949

THE FUNCTION OF EXCHANGE RATES. *By* R. G. HAWTREY . 145

A COMMENT. *By* SIR HUBERT HENDERSON 157

PROFITS IN BRITISH INDUSTRY FROM 1924 TO 1935. *By* RONALD HOPE 159

SOME THOUGHTS ON THE NATURE OF INTEREST. *By* E. VICTOR MORGAN 182

STATIC MODELS AND CURRENT PROBLEMS IN INTERNATIONAL ECONOMICS. *By* T. BALOGH 191

THE THEORY OF COMPARATIVE COSTS RECONSIDERED. *By* G. M. MEIER 199

OLIGOPOLY AND IMPERFECT COMPETITION. *By* E. G. DOWDELL 217

THE FOUNDATIONS OF WELFARE ECONOMICS. *By* I. M. D. LITTLE 227

THE ORIGINS OF INDUSTRIAL PEACE: THE CASE OF THE BRITISH BOOT AND SHOE INDUSTRY. *By* ELIZABETH BRUNNER 247

A NOTE ON CURRENT MARXIST DEFINITIONS OF THE NATIONAL INCOME. *By* DUDLEY SEERS 260

A NOTE ON PRICE AND QUALITY. *By* N. H. LEYLAND . . 269

Editorial Board

T. WILSON (*General Editor*) F. A. BURCHARDT
SIR HUBERT HENDERSON G. D. N. WORSWICK
P. W. S. ANDREWS (*Secretary*)

THIS SERIAL PUBLICATION is intended primarily as a channel for the publication of articles by Oxford authors. Contributions from elsewhere will, however, be considered, and should be sent to the General Editor. While most of the articles will be concerned with Economics, the editors also hope to publish occasional articles on Economic History, Public Administration, and Scientific Method, provided these are likely to be of general interest to economists.

THE FUNCTION OF EXCHANGE RATES

By SIR HUBERT HENDERSON

I

THE idea that variations in foreign exchange-rates can serve as an efficient regulator of a freely working international economic system has gained an astonishing vogue among British academic economists. It seems, indeed, to have become a main article of faith with those who still regard a return to a freely working international economic system as a righteous cause and a practicable goal. This, it is worth noting at the outset, implies a remarkable change of view. It is not very long since the extremists of economic liberalism were disposed to stress, and even to exaggerate, the advantages of fixed exchange-rates. In the era of international economic expansion before 1914 a rigid fixity of exchange rates was, of course, secured by the nearly world-wide operation of the gold standard. After the First World War, when exchange rates were in confusion and runaway inflation widespread in Europe, international economic statesmanship concentrated on the restoration of exchange stability through the medium of a general return to the gold standard, as a central aim of reconstruction. This policy commended itself to public opinion at large, and to the economic liberalism of the day in particular, for many reasons. Among them was the consideration, emphasized in the resolutions of international conferences, that the restoration of exchange stability was an essential pre-condition of the removal of import prohibitions and quota restrictions and other new varieties of trade impediments. During the 1930's, after the collapse of the reconstituted gold-standard system, the persistent refusal of the British authorities to re-establish a definite gold parity for sterling was criticized and censured in ultra-liberal circles as a regrettable piece of self-regarding economic nationalism. To economists like Professor Robbins and Sir Theodore Gregory, to bodies like the International Chamber of Commerce, fixed exchange rates were still a condition of international economic health.[1]

A Rip van Winkle familiar with the controversies of the last decade might, therefore, be startled if he were to wake up to-day to find that exchange-rate variations have been assigned a high constructive role in the ultra-liberal philosophy as the regulating spring of the future international mechanism, and the true alternative to the hated quantitative regulation. It may be instructive to consider how this change has come about. In the 1930's advocacy of fixed exchange parities implied acceptance of the whole gold-standard technique; and an essential part of that technique was that the amount of money and credit and purchasing-power within a country was related to, and governed by, the volume of its gold reserve. This relation provided the regulator needed for a self-adjusting mechanism. A large or persistent deficit in the balance of payments of a country would lead to an outflow of gold; this would enforce a curtailment of the volume of credit and purchasing-power, reduce incomes, prices, and costs, check imports and stimulate exports, and thus restore external equilibrium.

---

[1] 'The bitter experience of the World Economic Conference has shown that, unless the future of exchange rates is tolerably certain, it is quite hopeless to expect any substantial abolition of the obstacles to trade. You cannot negotiate about tariff rates effectively unless you have some idea of the future course of values. And you can have no idea of the future course of values unless you can rely upon a more or less steady rate of exchange. After all, many of the obstacles which it would be most desirable to remove, actually came into being because of the difficulties caused by fluctuating exchanges. Until the exchanges have ceased to fluctuate, it is surely futile to hope that they will be removed.' Professor Lionel Robbins, *Lloyds Bank Monthly Review*, April 1935.

Such at least was the picture of the chain of remedial reactions which time-honoured theoretical analysis had imprinted firmly. Early in the inter-war period heretics had begun to mutter that in practice the process was less simple and less satisfactory. An outflow or inflow of gold might arise not only from disequilibria in the balance of payments on current account but from capital movements; and capital movements might represent not the long-term international investment which had won a good name in the nineteenth century, but a scurry of individuals or businesses in other countries to put their money abroad for greater safety or to bring it back again when their confidence recovered. Such erratic movements of 'hot' money were not a good reason for making credit in Great Britain abnormally easy at one moment and tight at another. Yet this was the natural outcome of the automatism of the gold standard.

After the fall of the pound in 1931, this consideration told strongly against returning to gold. By this time, moreover, 'hot' money movements had attained such large dimensions as to convince our authorities that a reserve of international purchasing-power much larger than the traditional gold reserve of the Bank of England had become indispensable. The Exchange Equalization Fund was accordingly created in 1932; and as an incident of this, the link was severed which had hitherto connected the amount of our internal purchasing-power with the monetary reserves available for external payment. It thus became possible to insulate our money market and credit system from the disturbing effects of quite large movements of funds whether in or out. Our monetary authorities came increasingly to appreciate how convenient this was for various purposes, notably for the uninterrupted pursuit of the policy of cheap money designed for the promotion of trade recovery. In this way a sense of the increasing disadvantages of the automatism of the gold standard under inter-war conditions gradually merged into a cautious sense of the positive advantages of a more insulated system in opening out new possibilities of deliberate and constructive policy.

The extremists of international economic liberalism, however, viewed these developments with disapproving eyes. They maintained that the firm re-establishment of definite exchange parities was the true corrective for 'hot' money movements, and that, for this, a readiness to stabilize sterling was indispensable. They were not attracted by the greater scope for deliberate policy under an autonomous monetary system, for they were predominantly *laisser-faire* in outlook. The more reflective of them expressly urged that the link between internal purchasing-power and reserves of gold and foreign exchange should be restored.[1] Thus the free international economic mechanism which they wished to reconstitute would have contained its traditional regulator.

This is no longer true. Apparently there is no one left to-day who desires to make the volume of our internal purchasing-power vary with the size of our monetary reserves. For the future, it is to be determined, according to the now fashionable prescription, solely with a view to maintaining full employment. Subject to some cautionary words about this phrase, Professor Robbins has declared his adherence to the policy. 'I favour something', he wrote in 1947,[2] 'which, if you like, you can call over-all financial planning' designed 'to maintain over-all stability of aggregate demand, while leaving the maximum flexibility between the various constituent items.' This, it should be observed, is recommended not for a temporary emergency but for the period 'when the problems of the transition are over'. Professor Robbins recognizes that it implies a certain change of view ('I confess that I have not always held this conviction as strongly as I do to-day'); but the change is far more radical than these words suggest. It means shifting the criterion by which the effective volume of our purchasing-power is to be determined from one related to our international position to one related exclusively to our internal conditions. This would certainly have aroused Professor Robbins's strong disapproval before the war.

To all appearances this change of view is widespread. A purely internal criterion for regulating the volume of purchasing-power is accepted to-day by those who still hope for a return to a free international economy no less readily than it is by others. Faith in the old regulator of the gold standard seems to have faded quietly away during the war years. Perhaps the appearances are deceptive; it may be that there are some who have not altered their views, but are merely silent and perplexed. Certainly it is somewhat curious that the war should have effected this revolution of opinion. For it would be possible to argue, not unreasonably, that the gold standard might work more satisfactorily in future, now that exchange control over capital movements has been introduced in Great Britain and many other countries. Not that this control is ever likely to be so effective as to prevent a slow, steady, capital leakage, if the underlying economic forces are set strongly in that direction. But it should suffice to avert large, sudden, fortuitous movements of 'hot' money. Accordingly changes in a country's monetary reserves may be expected to reflect more accurately than in the past the state of its balance of payments on current account, and are only likely to be modified or aggravated by a strong persistent tendency towards an outflow or inflow of capital, to which it can be argued that internal credit conditions *ought* to be responsive. On these grounds, the line might well have been taken that the gold standard would be sufficiently modernized by retaining exchange control over capital movements, and that it should be an aim of policy to restore its essential elements, namely the fixity of exchange rates and a definite relation between internal purchasing-power and monetary reserves.

---

[1] 'Fluctuations in the resources of the Exchange Equalisation Fund should be treated in the same way as fluctuations in the gold reserve in a centre finally stabilised on gold. If the reserve shows signs of serious depletion, then the rate of discount should be raised and steps should be taken to make it effective. No doubt this sounds disagreeable. We have become so accustomed to the drug of ultra-cheap money, that the thought that, to enjoy the benefits of international revival, we may be compelled to forgo the benefits of ignoring international conditions is very irksome. We like to think that international revival will be super-added to the delights of irresponsibility. We do not see that the failure of international revival to arrive is in part at least the price for the policy we have hitherto pursued.' Professor Robbins, *Lloyds Bank Monthly Review*, April 1935.

[2] *The Economic Problem in Peace and War.*

In fact, however, the line taken has been quite different. During the war-time phase of post-war economic planning it became customary to do obeisance to a somewhat arbitrary potentate, known as 'the climate of opinion'; and on nothing, it was understood, was this potentate so insistent as on ensuring full employment and avoiding deflation in any circumstances. In this atmosphere the extremists of economic liberalism were chiefly concerned to argue that the object could be attained quite simply and without any serious interference with a free economy, by the expedient which Professor Robbins calls 'over-all financial planning'. This contention was rendered more plausible by a new, ingenious argumentative technique, invented, I think, by Lord Beveridge and widely adopted since. Awkward difficulties, such as the repercussions of a strong demand on wage-rates and prices, which would previously have been recognized as belonging to the essence of the problem, were disposed of by recognizing them, with apparent candour, 'assuming' that they would be overcome somehow, and then ignoring them. With the aid of this device many economists have been able to hypnotize both themselves and the more sophisticated sections of public opinion into the remarkable belief that the problem of averting serious unemployment in future has been solved triumphantly by modern abstract economic analysis, unsupported though this is by experience, and resting though it does on the manifestly unreal hypothesis of a 'closed economy'.

They have been unable, however, to blind themselves to the fact that the new 'over-all financial planning' must involve scrapping the old regulator by which maladjustments in the balance of international payments were corrected. The idea of deliberate regulation by such means as import restrictions and quantitative programmes is peculiarly repellent to all price-system ideologues, except as a strictly temporary expedient. For normal conditions they feel it essential to find a new automatic regulator; yet none appears to be available, other than exchange-rate variations. An optimistic view of their efficacy has become accordingly for many a necessity of the logic of their position. This goes far, no doubt, to explain the incomplete thinking and lack of critical sense and common sense with which exchange-rate variations are commended.

II

Let us consider what has to be assumed to justify the belief that exchange-rate variations might serve as an efficient regulator. First, it is clearly necessary to assume that a depreciation of a country's exchange rates will normally improve its balance of payments, and that this effect will be a lasting one. For this to be true an indispensable condition is that the level of internal prices and incomes should not rise proportionately to the depreciation even in the long run. On this hypothesis it is common ground that the effect of the depreciation on the balance of payments will be the resultant of opposing tendencies: on the one hand a worsening of the country's terms of trade, on the other hand an increase in exports and a decrease in imports in terms of volume. Therefore the further assumption is required that the latter tendency will preponderate over the former.

The reasonableness of this last assumption has been the subject of much discussion and controversy, necessarily inconclusive. The essential truth about it can be summed up, in my opinion, in the following propositions:

1. It is normally helpful in the long run to a country's balance of payments that its currency should be 'under-valued' rather than 'over-valued'; that is to say, that its internal value, representing its power to purchase labour and home-produced goods, should be somewhat higher than its external value, representing its power to purchase imported goods, and depending partly on exchange rates and partly on the price-level in the outside world. The presumption that there will be a net advantage arises partly from the fact that an under-valuation involves a fall in the aggregate of national income and

purchasing-power, when reckoned in terms of other currencies. The effect on the volume of imports and exports tends, therefore, to be greater than that which would be produced by price influences alone.

2. Nevertheless, this net advantage will not increase indefinitely with the extent of the under-valuation. After a certain point the unfavourable effects of a further under-valuation on the terms of trade will outweigh the favourable effects on the volume of imports and exports. This certainly is the conclusion suggested by the experience of run-away inflations, during which the extent of the under-valuation existing at any moment was usually extreme.
3. For correcting an adverse balance of payments there is, therefore, in principle, an optimum degree of under-valuation; and this, though it may vary from one country to another and from time to time, is usually, in my judgement, slight. I use the words 'in principle' because I do not wish to suggest that either the actual degree of a currency's under-valuation, or the degree that would be most desirable, can be precisely calculated.

I have dealt somewhat cursorily with this question because I do not regard it as the really crucial one. It seems to me more important to examine closely the questions that arise directly out of it. How is the optimum degree of under-valuation, whatever it may be, to be established and maintained? Is it likely, or even conceivable, that the expedient of exchange-rate adjustments could serve this purpose effectively? In considering these questions we must never lose sight of the fact that the under-valuation which has to be established is not a simple entity but a tripartite relation, in which rates of exchange are only one factor, the other two being the level of costs and prices in the country concerned and the level of prices in the outside world.

The effect of a depreciation of exchange rates on the above relation is likely to depend in practice on what is happening to the last factor, which, of course, lies outside the control of national policy. It can reasonably be expected to be a beneficial one, from the standpoint of the balance of payments, as from others, during a severe world depression, when prices in terms of gold and other currencies are falling sharply. In these conditions an exchange depreciation, within due limits, will not serve to raise internal costs and prices, except possibly by way of recovery from a very recent fall. The broad effect will be to avert or mitigate the need for a fall in wage-rates and internal prices in order to maintain the country's competitive position in world markets. Now the process of adjustment to the need for a lower price-level is notoriously slow and painful, and is apt to be incomplete. Wage reductions are never easy to bring about. Some prices and money incomes are more or less 'rigid' for long periods, or are at least 'sheltered' from the pressure of international competition. Thus the real incomes of many individuals tend, under deflationary conditions, to increase. This throws an additional strain on others; and the less sheltered producers must suffer a reduction of real income, as well as money income, if the country's competitive position is not to be impaired. To be relieved from the need to effect a readjustment of this sort, and from the pressure towards it, is a real service and may be a crowning mercy.

It was in conditions of severe world depression and deflation that the fall of the pound in 1931 proved of great advantage to the British economy, and helped for a time to improve the competitive position of British industries in world markets and the British balance of payments. The idea that exchange depreciation, previously regarded as a calamity, might rather be a blessing came into fashion as the result of this experience, and received an early and startling expression in the deliberate forcing down of the gold value of the dollar in 1933. The events that followed showed that this idea had serious limitations and drawbacks even in a deflationary setting. The depreciation of the pound and the dollar had disastrous repercussions on the economies of the gold *bloc* countries, notably France; which eventually were forced to devalue too; and there seemed at one time a real danger of a chaotic process of competitive exchange depreciation, which was averted by a 'gentlemen's agreement' between the United States, Great Britain, and France to avoid deliberate exchange depreciation for the future, known as the Tripartite Agreement of 1936. This experience does nothing to encourage the view that exchange-rate adjustments by particular countries would form a good basis for an orderly international economy. It remains true, however, that if serious world depression and deflation should recur, exchange depreciation would be useful and perhaps indispensable to countries in balance-of-payments difficulties.

It cannot be inferred that it would be equally helpful when world prices are rising, or even when they are more or less steady. In such conditions a lowering of exchange rates would cause internal prices to rise absolutely, and not merely relatively to prices in the outside world. The rise would not be confined to the prices of imported goods, but would spread gradually throughout the country's economy. In Great Britain more especially, an exchange depreciation would contribute either to a wage-inflation, by raising the cost of living, or to a financial inflation, if the attempt were made to counteract this by additional food subsidies. In any other environment than that of world deflation, a substantial exchange depreciation must be a strong price-raising influence.

If it were not for the dynamic quality of inflationary movements, it would no doubt be arguable that the rise in internal prices and costs ought to be less than proportionate to the exchange depreciation; and attempts were, I think, made at one time to calculate the degree of inflation that would represent an appropriate adjustment on the basis of an assumed rigidity of certain prices and incomes. In practice, however, the extent of an inflationary movement, once it has been set going, cannot be forecast. The process is apt to be self-perpetuating, and indeed to gather momentum, unless and until it is checked by some obstacle or counter-force, for which there is no provision in the technique of exchange depreciation. There can be no assurance therefore that the rise of internal prices would stop at any theoretically appropriate points. It would be a happy accident if it did.

This raises an awkward question, to which it is important to know the answer. What is supposed to happen, under a system of exchange-rate variations, if the rise of internal costs and prices following a lowering of exchange rates should overshoot the mark, and instead of leaving some degree of under-valuation, helpful to the balance of payments, should result in an over-valuation ? So far as I know the advocates of variable exchange-rates are silent on this point. But clearly there are only two courses open. One is to lower the exchange rates again in the hope that there would be better luck next time. If this course were followed there would be no safeguard against the danger of a vicious spiral of exchange depreciation and internal inflation, gathering speed and continuing indefinitely. The other course is to call in aid whatever instruments may be most effective for forcing prices down, including perhaps the restriction of credit. These are the instruments which were brought into play more or less automatically under the gold standard, when gold was flowing out. If they have also to be used after an exchange depreciation, a further question arises. What would have been gained by the preliminary depreciation ? If it may be necessary, after exchange rates have been lowered, to use the traditional deflationary instruments to establish a certain degree of under-valuation, why not use them in the first instance to establish the under-valuation at the original parities ?

Reasons can be urged, it is true, for preferring an initial devaluation, which sometimes possess great cogency. For example, the real value of the interest payable on the national debt would be reduced by an inflation; and when the proportions of the debt charge to the national income and the public revenue are extremely high, it may be important that the public finances should obtain some such relief. Upon this issue much, of course, can be said on the other side; and I do not personally believe that the case for devaluation as a debt-relieving measure is nearly made out in Great Britain. But the point to note is that arguments of this sort for a once-for-all devaluation provide no support for the idea of using variations in exchange rates as a balance-of-payments regulator. If this idea were really to be applied, and if an exchange depreciation were followed by a fully proportionate inflation, there would be no escape from the other horn of the dilemma, namely, a further lowering of exchange rates and the risk of a vicious spiral, in which the internal value and the external value of the currency would chase each other downwards.

The fundamental impracticability of the idea now begins to emerge; and, in order to bring it out clearly, it may be helpful and appropriate to use language of quasi-mathematical formalism. Price-system forces can only serve to regulate or improve the balance of payments by establishing and maintaining a favourable relation between three factors: exchange rates, internal prices, and world prices. One of these, world prices, is outside the control of national policy. Therefore it must be possible to cause one of the other two factors to move in whichever direction is required, while the third is held firmly at a fixed point, or at least is effectively prevented from changing in equal proportion. It is possible to cause internal prices to fall, if that is the movement required, by the use of deflationary monetary and financial instruments, while the exchange rates are held firmly at fixed parities; though there are limits to the extent to which this process can be carried in practice. This was the *modus operandi* of the gold standard. It is *not* possible to reverse the roles; i.e. to lower the exchange rates, and to keep internal prices stable or to ensure that they will not rise beyond a certain point. For the deflationary, or counter-inflationary, instruments are not precision instruments. If applied with sufficient force and for a sufficient length of time, they are capable of reversing an inflationary trend; they are incapable of the delicate work of allowing prices to rise so far and no farther. Still less could they hold prices and money incomes at a constant level, if exchange rates vary.

This would be a fatal weakness of the method of exchange depreciation, even if the regulation of internal prices were made the governing purpose of monetary and financial policy. In fact, however, this is not recommended, at least not ostensibly, by the advocates of exchange-rate variations. As I have indicated above, they subscribe, with whatever qualifying hesitations, to the doctrine that these instruments should be used to maintain aggregate demand at a level conducive to full employment. It is important to explore somewhat further the way in which this doctrine complicates the problem.

III

Nothing has done more in my opinion to lead modern economists astray than the practice which Lord Keynes brought into fashion in the 1930's of reasoning on the hypothesis of a 'closed economy'. Elaborate structures of theoretical analysis, purporting to supply light and guidance for the central economic problems of the actual world, have been built on what is in effect the assumption that there is no such thing as international trade, that all economic transactions take place within the boundaries of a single country and are therefore within the scope of its Government's control. This assumption is, of course, supposed to be merely a provisional simplifying hypothesis, introduced to facilitate the orderly development of thought; and in principle it is recognized that all conclusions reached upon this basis must be subjected to subsequent revision in the light of the fact that our actual economies are far from being 'closed'. Unfortunately this final task is apt to be very perfunctorily performed and sometimes neglected altogether.

It is paradoxical that British economists should have led this intellectual fashion, at which Adam Smith, Ricardo, John Stuart Mill, Bagehot, and Alfred Marshall would certainly have knit their brows. International trade has been for centuries a factor of fundamental importance in the British economy. Even before 1914, our most serious economic troubles had their origins abroad; fluctuations in the overseas demand for British exports, both of consumers' goods and of capital goods, did far more to influence the level of our industrial activity than fluctuations in, say, the home demand for house-building, which frequently moved in the opposite direction from the general trend of trade. At any time, therefore, within the modern era, it would have been inappropriate and misleading to treat British economic problems as though international developments were a mere complication which could be ignored in a provisional analysis. That this fashion should have sprung up after the 1914–18 War is still more paradoxical. To-day our main economic problems and difficulties are clearly those of international trade and international payment. So they were, in a somewhat different form, throughout the inter-war period. In such conditions, generalizations and prescriptions of policy, suggested by theoretical analysis proceeding on the hypothesis of a 'closed economy', are apt to be dangerously out of focus.

Yet it is on this basis, and with a complete disregard of this danger, that fashionable economic analysis has pursued its examination of the problem of unemployment, and claims to have 'discovered' its secret. This claim was readily accepted at its face value by a wishful-thinking public during the war years, and received some measure of official endorsement in the White Paper on Employment Policy issued by the National Government towards the end of the war. The key to the problem, according to this creed, is to maintain aggregate demand at an adequate level. Therefore, if a substantial contraction of demand should occur in some sectors of the economy, it is proposed that a compensating expansion of demand should be elicited in other sectors, by whatever means may be most efficacious for the purpose.

As a general proposition this seems attractive; but it includes the important special case that if our export sales were to diminish, so that the demand for labour in the exporting industries were to be reduced, a compensating increase of demand must be stimulated in the home market. Clearly this implies a passive acceptance of the worsening of the balance of payments resulting from the loss of export trade; and the policy contemplated would worsen it still farther. The expansion of home demand would lead directly to an increased demand for imported goods. If it were to be large enough to overcome the serious 'structural' obstacles to a 'transfer' of labour, it would involve an increase in wage-rates and costs of production detrimental to our competitive position in world markets. Furthermore, in so far as work-people displaced from the exporting industries were successfully absorbed in other industries, it would become more difficult for the former to re-acquire a sufficient labour force if the demand for their products were later to revive.

As the doctrine of maintaining full employment through the instrumentality of a sufficient aggregate demand was evolved in a 'closed economy' setting, these difficulties were kept well in the background. Moreover, when attention was called to them, it could be argued in the 1930's, without evident absurdity, that they constituted no more than a minor drawback, and that the maintenance of full employment was greatly preferable to a better balance of payments. To-day, however, it should be manifest that the complacent indifference to repercussions on the balance of payments which the policy presupposes is out of place. In the years that lie before us we shall be in no position to adapt our economy to a loss of export trade by measures which would make the balance of payments worse.

The policy in question would serve directly, in the circumstances in which it would be most likely to have to operate, to enlarge a balance-of-payments deficit. But this is not all. It would rule out the possibility of correcting the deficit by restricting the volume of internal purchasing-power. It would therefore deprive us, if the instruments of quantitative regulation and direct control were also to be laid aside, of any effective means of maintaining exchange stability, if that were the *desideratum*. This again is not all. The policy would be fatal to any hope of correcting the deficit by the method of exchange depreciation, slender as this hope would be in any case, if my argument in the preceding section is correct. For it would be impossible to apply the instrument of a limitation of internal purchasing-power to the difficult task of preventing the rise in internal prices and costs from outstripping the exchange depreciation.

For any country in a weak balance-of-payments position the deadlock seems complete. Those who have concluded so hastily that the secret of maintaining full employment has been discovered in the formula of a sufficient aggregate demand must think again.

It is fair to say that the more recent statements of this formula reveal misgivings as to some of its implications. But these misgivings seem at present to be limited to the implications that are relevant to issues of internal policy. Professor Meade's recent book, *Planning and the Price Mechanism*, provides a curious example of this limitation. He is perturbed at the possibility that a high aggregate demand may lead to excessive wage-rates, and he has a particular dislike for the practice of raising wages to meet increases in the cost of living. He refers to the danger that this practice may lead to a loss of export trade; and to this warning he attaches a footnote (p. 74) which is worth reproducing in full:

> 'Unless, of course, we offset the rise in our internal costs by a depreciation of the exchange rate, in which case the price of imported foodstuffs and so the cost of living would rise still further. If this in turn were followed by a rise in wage-rates, we should again either lose export markets or have to depreciate sterling still further, which would again raise the cost of living. And so on *ad infinitum*.'

It seems reasonable to infer that exchange depreciation would be a most dangerous remedy for a balance-of-payments deficit. Yet, in his final chapter, Professor Meade advocates this expedient for long-run use with ardour, untroubled by doubt. He argues at length its superiority, as a price-system remedy, over import restrictions, ignoring altogether the possibility that incomes and prices might rise. Indeed he goes so far as to assert (p. 96) that 'the case for exchange rate adjustment is conclusive' if 'a relatively small reduction in the price of our exports in foreign markets [would] cause a relatively large increase in the volume of goods which we could sell'. This is a palpable *non-sequitur*, if it is admitted that, owing to wage increases, there might be no lasting reduction in the export prices we were able to quote.

Professor Meade may perhaps suppose that, having urged that wages should not rise because the cost of living rises, he is entitled to assume that they will not do so. He does not, however, assume the difficulty away so wishfully in discussing questions of internal policy. He lays stress (p. 71) on the 'most bitter dilemma' with which society may be faced, if trade unions insist on exploiting their bargaining strength under high employment, in having to choose between continuous inflation and substantial unemployment. Professor Robbins in *The Economic Problem in Peace and War* discusses the same dilemma along very similar lines, and offers the formula that the aim should be 'to maintain aggregate demand at a level which, *at current or slightly rising rates of wages*, would secure a reasonably high level of employment and utilization of resources' (p. 71). (The italics are his own.) But he gives no hint that the dilemma has any bearing on our external economic problem.

Two crucial questions lie concealed within the folds of Professor Robbins's formula. Does he, too, mean that money wage-rates should remain unchanged, or rise only slightly, whatever may be happening to the cost of living? And does he mean that the volume of purchasing-power should be so regulated as to prevent wage-rates from rising too much? If the answers to these questions are in the affirmative, it is evident that Professor Robbins is expecting a great deal from the financial and monetary instruments available. To prevent substantial wage increases following a sharp increase in the cost of living, it is probable that a drastic restriction of credit, involving heavy unemployment, would be required. A policy of this sort would differ radically from anything that has hitherto been thought of as a policy of full employment, however qualified. Under it, the criterion for regulating credit and purchasing-power would have been shifted in effect from the state of employment to the movement of wage-rates.

This suggests a further question. If the criterion for regulating purchasing-power must be shifted in practice from the fair-seeming but deceptive one of the state of employment, with no further questions asked, might not the most satisfactory alternative in some conditions be the time-honoured criterion of the size of the monetary reserves, suitably adapted to modern conditions? No country, after all, can afford to permit its monetary reserves to disappear. If they are dwindling fast, action has got to be taken to arrest the decline; and it is important that this action should be timely and effective. Whatever general principles may be proclaimed, an undue loss of monetary reserves must, therefore, be a cogent and compelling reason for restricting the volume of purchasing-power. It seems to me of real importance that the proclaimed principles should be revised so as to recognize this fact expressly. A subordinate issue also arises. Governments might take appropriate action to limit purchasing-power from their own sense of the need for it, without being required to take it by statute. Indeed our main hope must be that they would do so. But there is always the danger that wishful thinking, indecision, obscurantist prejudice, political timidity or faction might cause them to do too little and too late. I agree with Professor Robbins in thinking it unwise on these vital matters to trust entirely to governmental wisdom and initiative, and in desiring to supplement them by formal regulations prescribing the action to be taken in defined circumstances, so long as these 'automatic' arrangements are well designed.

I suggest that it would be wise in Great Britain to provide that, if and when our monetary reserves fall below a defined figure, chosen as the minimum compatible with safety, a substantial reduction of the note issue must follow automatically, and that this arrangement should only be alterable by a new Act of Parliament. A reduction of the note issue would, of course, carry with it higher short-term interest-rates and a restricted volume of bank credit. There are reasons, which in ordinary circumstances are probably good, for thinking this a clumsy method of curtailing aggregate demand, and for preferring others such as an over-balanced budget and the limitation of long-term investment work. Therefore, so long as the reserves are above the specified minimum, the discretion of the Government to use whatever means they think best for regulating aggregate demand should not, in my view, be fettered by automatic currency arrangements. Also Governments should not be hampered by misconceived international obligations, or by ideological prejudices, in using the more direct methods of import regulation and trade policy to keep the reserves from falling to the danger point. But when this point is reached, other methods must be used as well. The method of credit restriction, clumsy though it may be, is one of proved, if limited, effectiveness; it would be essential to fall back on it as a weapon of last resort, and prudent to provide in advance that this should be done in time.[1]

This, then, is the basic principle which I suggest for the future regulation of the volume of purchasing power: above a specified minimum of gold and hard currency reserves, unfettered discretion for the authorities, who should, however, look to the movement of the reserves as a main criterion of policy; below this minimum, an automatic and substantial curtailment of currency and credit. In putting forward this principle for consideration, I am not thinking only of Great Britain, though I have stated it for the sake of clarity in British terms. It might be of greater immediate importance to secure its effective application in other countries, e.g. France.

IV

The two essential principles of the gold standard were fixity of exchange rates, and a definite relation between internal purchasing-power and monetary reserves. I have just suggested that we should do wisely to reincorporate the latter principle in our monetary arrangements, though in a greatly modified form. It will be evident from the general tenor of my argument that I also think it would be wise to reintroduce the former principle, though again with modifications.

---

[1] Since the arrangement would be designed to operate only if the situation had become serious enough to require really drastic action, the minimum figure might well be considerably less than the £500 millions, the maintenance of which has been declared to be an object of 'unalterable policy'. It might perhaps be put as low as £350 millions. On the other hand, it would be important to ensure that a fall of the reserves below this minimum would entail a contraction of currency and credit severe enough to produce results. The regulations should therefore be so contrived as to enforce a reduction of the note issue by some appropriate multiple of any short-fall of the reserves below the minimum, say three times.

The idea that exchange-rate variations might be used systematically to correct maladjustments in the balance of payments is not only, as I have argued above, fundamentally impracticable. It misconceives the attributes of exchange rates; it overlooks their special fitness for a function of real importance. They are one of the few things in the economy which can be definitely fixed; and it is well that something should be fixed. The true function of exchange rates is to provide a constant factor round which the more variable elements may move, and by reference to which they can be adjusted; in other words, to provide a focus of stability in an orderly price system. The confidence which was felt in the fixity of exchange rates under the gold standard contributed much to the remarkable development of international economic life during the nineteenth century. It also contributed more generally to the comparative efficiency with which an unregulated price system then worked. Something valuable is lost when this confidence has disappeared. The value may have been overrated formerly; to-day it is certainly underrated. In Great Britain the force of this general consideration is greatly increased by the special considerations relevant to her position as the centre of the sterling area.

The modifications which are desirable under modern conditions can best be judged by reviewing the interrelated defects and mistakes that were responsible for the breakdown of the attempt to restore the gold standard as an international system after the 1914–18 War. Prominent among these was the instability in the purchasing-power of gold, expressed in the heavy fall in commodity prices during the world crisis of 1929–33. If in the future severe world depression and deflation should recur, any arrangement designed to ensure stable exchange rates would again break down. In such conditions, as I have tried to show above, individual countries can derive a much-needed relief from exchange depreciation, and the use of this expedient might become imperative in countries with balance of payments difficulties. For different and more complex reasons, a continuing inflation of world prices would also prove incompatible with fixed exchange rates. Therefore, a reasonable stability in the purchasing-power of gold, if that is to remain the standard in terms of which exchange parities are defined, is an essential condition of international exchange-rate stability.

In practice the purchasing-power of gold will depend mainly on the policies pursued by those countries which are untroubled by balance-of-payments anxieties and equipped with abundant gold reserves. Therefore, a special responsibility for creating the conditions in which exchange-rate stability can be maintained rests upon these countries, outstanding among which is, of course, the United States. In the meantime, uncertainty upon this matter would suffice to forbid the idea of attaching to modern exchange parities the sacro-sanctity which they possessed in the hey-day of the gold standard. The possibility of modifying these parities, as a *pis aller*, must be left open, though it must not be supposed that a new and more orderly international system can be built upon this basis.

Another main cause of the breakdown of the gold standard in 1929–33 was the excessive strain which was thrown upon its mechanism by large maladjustments in the balance of international payments. Since the maladjustments that have now to be corrected are far larger, it is important to interpret rightly the lessons of this experience. No blame can fairly be attached to the principle of exchange-rate fixity. Some responsibility is attributable to the fact that the exchange rates for sterling were overhigh as the result of the decision of the British authorities to return to gold at the traditional parity. But from the international standpoint the chief mistake that was made, in my opinion, was to entrust the task of restoring equilibrium far too exclusively to the forces of the price system. Half-hearted attempts to assist these forces by deliberate policy and measures of direct control were deprecated and discouraged as contrary to sound principle and obstructive to international trade. Countries in balance-of-payments difficulties were persuaded to rely instead on the aid of foreign loans, floated in the centres of the more strongly placed countries, and organized and developed on a large scale with the help of the goodwill and public-spirited exertions of the dominating figures in the banking world and the officers of the League of Nations. The tradition of economic liberalism was then too strong and self-confident for any other type of policy to have been feasible. The event proved, however, that the aid given by private international lending was unreliable and fitful, and that the more automatic forces of the price system were too weak to do the work expected of them.

For the much larger task of readjustment which faces us to-day, deliberate policy, using such instruments as systematic import programmes and trade arrangements, is not only indispensable for the time being, but will remain indispensable in my judgement for many years to come. To suggest, as is sometimes done, that, though import restrictions are necessary now, they may become superfluous comparatively soon if an improved price system is given freer play, is to nurse illusions. As I have argued elsewhere,[1] the British balance-of-payments deficit could only be corrected by price-system forces at the cost of a huge, unnecessary, and impracticable lowering of the standard of life.

A further cause, and superficially the most important, of the 1929–33

[1] Address delivered at the British Association. *Economic Journal*, Dec. 1948.

breakdown was the overwhelming magnitude of the short-term capital movements which developed. Despite their imperfections, the systems of exchange control which have been introduced in many countries should provide, it may be hoped, a fairly effective safeguard against a recurrence of this particular disorder. So far as can be foreseen, the need to retain exchange control for the purpose of limiting capital transfers will continue indefinitely. This represents, therefore, an essential change in the arrangements by which fixed exchange parities used to be maintained.

In ways like these it is important to adapt our monetary arrangements to the altered conditions of the difficult modern age. It is a profound mistake, however, to suppose that it would be wise or helpful to invert the principles which underlay the traditional monetary scheme, by making variable that which was formerly fixed, and by trying vainly to fix that which was formerly variable. The economic failure of the inter-war period was due less to the imperfections of the monetary system than to the policy which left it, insufficiently supported, to perform a task of readjustment which was beyond its powers, and would have been equally beyond the powers of any other arrangements of a price-system type. Exchange-rate fixity, like other desirable things, may prove unviable in times of extreme disorder; but it is a most desirable thing, and it should be an important aim of policy to establish and maintain it.

ALL SOULS COLLEGE,
OXFORD.

NOTES ON BRITAIN'S BARGAINING POWER

By G. D. A. MACDOUGALL

1. Introduction

BRITAIN has hitherto been the largest importer in the world. Table I shows that, in 1937, her imports were 17 per cent. of the world total, and that she took no less than 30 per cent. of the world's imports of food. Her share in world imports of certain commodities was even larger—four-fifths of the world's imports of meat and butter and half the world's imports of eggs, cheese, and tea. Certain countries, too, depended greatly on the United Kingdom market. More than half of Denmark's exports came to this country, over four-fifths of New Zealand's, and over nine-tenths of the exports of Eire.

TABLE I

United Kingdom and World Imports 1937

	U.K. imports (*Billion dollars*)	*World imports* (*Billion dollars*)	*U.K. as per cent. of world imports*
Food	1·9	6·4	30
Raw materials . .	2·0	11·6	17
Manufactures . .	0·8	10·2	8
	4·7	28·2	17

The British market is thus of vital importance to many primary producers throughout the world; and it is often argued that, by threatening to reduce, to eliminate, or not to increase our imports from particular countries we might induce them to help us pay for our essential imports. We might, for example, induce them to charge us a lower price for their produce or to pay a higher price for our exports; to take more of our exports, either in addition to those they were already taking from other countries or in their place; to reduce their import barriers against manufactures from all countries; or to assist us in other ways.

It is clearly of some importance to evaluate, in more detail, our bargaining power as a large importer, and to consider the possible repercussions of the use of this power as part of our commercial policy.

In the *Economic Journal* for March 1946 I published a brief analysis.[1] This seemed to show that many countries were less dependent on the British market than was commonly supposed, and that many were even more dependent on the market of the United States. The measure I then used of a country's dependence on our market was the proportion of its total exports sent to the United Kingdom. But, as I explained, this is by no means the only measure. The main object of these further notes is to consider certain others. Thus, for example, a more useful measure may sometimes be the proportion, not of a country's exports, but of its total output, sent to the United Kingdom. We shall also make an analysis commodity by commodity; for it has been argued that my analysis by countries understates the bargaining power of Britain.

---

[1] 'Britain's Bargaining Power.' A later article (*Economic Journal*, March 1947) included a further analysis of the bargaining power of the United States.

This article is by no means exhaustive. Many other considerations, which are not treated here, are highly relevant. We shall have nothing to say, for example, on the stability of our market as compared with that of other countries, or on the relative ability of this and other countries to offer loans and gifts, or to offer scarce commodities—a bargaining weapon of much potency in recent years. Nor shall we discuss the lessons to be learned from recent exercise of our bargaining power.

The figures relate to pre-war years throughout and are thus some 10 years out of date. This clearly needs some explanation. One reason is, of course, that many of the statistics required are not readily available for 1946 and 1947. But even if they were I should be reluctant to use them. It is doubtful whether the pattern of world trade in 1946–7 was any nearer a post-war norm than the pattern existing in 1937–8. Even more important, I believe that figures for 1946–7 would show Britain's bargaining power, according to the measures I have chosen, to be less than it was in 1938. British imports were only 13 per cent. of world imports in the later years compared with 18 per cent. in 1938 (while United States' imports rose from 9 per cent. of the world total in 1938 to 13 per cent. in 1946–7).[1] The exports of many countries have, moreover, recovered less quickly than their production, and have been a smaller proportion of their output than before the war. The proportion taken by the United Kingdom of the exports, and of the output, of many countries and commodities must thus have been lower in 1946–7 than before the war. But, as 1946 and 1947 were years of transition, it might be thought that figures relating to them would underestimate Britain's bargaining power in the future, and so unduly strengthen the general tenor of my argument. I therefore prefer to use pre-war figures for fear of giving a false general impression, although the details might be very different had more recent figures been used. I shall return in section 7 to the comparison of our pre-war and post-war bargaining power.

One final introductory remark. This paper is intended to give a general impression. The method is to present tables covering a fair number of countries and commodities but to avoid excessive detail and refinement. This means that some of the figures may be misleading if taken in isolation. It may be misleading, for many purposes, to combine the figures for mechanical and chemical wood-pulp, or for all types of cotton, in Table VI. Some of the national income figures on which Table V is based may be too high, others too low. When a table refers to a single year, some individual figures may be freaks. Those with expert knowledge of particular countries and commodities will find some of my figures hard to swallow. I freely admit in advance to any technical howlers they may discover. But I have a faith in the 'law of compensating errors', and I do not think the general picture will be affected. Clearly, however, no significance should be attached to the precise figures given. These depend on the choice of years and the method of calculation. I have given the sources from which the figures have been derived, and the reader may refer to these for definitions and qualifications. To mention them all would double the length of this paper.

---

[1] Calculated from the International Monetary Fund's *International Financial Statistics* for July 1948. The figures for the United States and for the United Kingdom refer to general trade. Elsewhere I have used figures for special trade.

2. Individual Countries' Exports to the United Kingdom as a Proportion of their Total Exports

Before considering other measures of our bargaining power it will be convenient to recapitulate the results of my earlier analysis and to suggest some morals that can be drawn from them.

Table II shows the proportion of their total exports sent to the United Kingdom in 1938 by the various countries of the world. (The gold exports of the gold-producing countries—or their output if lower—have been included in their total exports but not in their exports to the United Kingdom since, even when their gold was in fact shipped to this country, it is clear that they did not depend on us for its sale.)

Ninety-eight countries are distinguished in Table II (including a number of 'composite countries'). Of these, two countries (Eire and New Zealand) were almost wholly dependent on our market; both sent over 80 per cent. of their exports to us. Eight others depended on us for more than half their exports, but of these the only important trading country was Denmark, which sent 56 per cent. of its exports to the United Kingdom.

Next we find twenty-seven countries that sent between one-fifth and one-half of their exports to this country, the most important being India, Canada, and Australia; the Argentine; Norway, Sweden, and the Baltic States.

Finally, we find sixty-three countries (two-thirds of the total) dependent on us for one-fifth or less of their exports. These include most European countries south of the Baltic; the United States; most of tropical America; most of the Far East; and South Africa. The sixty-three include, it is

TABLE II[*]

Percentage of overseas country's exports[†] sent to U.K. in 1938	*No. of countries*	*Percentage of all U.K. exports in 1938 sent to each category of country*	*Main countries included. (Figures show percentage of all U.K. exports sent to each country in 1938)*
91–100	1	4·4	Eire 4·4
81–90	1	4·1	New Zealand 4·1
71–80	1	0·2	—
61–70	1	—	—
51–60	6	5·1	Denmark 3·4
41–50	6	11·4	Australia 8·2 Finland 1·2
31–40	11	20·6	India and Burma 7·9 Canada 4·9 Argentina 4·2 Egypt 1·9
21–30	10	11·0	Netherlands 2·8 Sweden 2·5 Norway 1·7 Iran 1·2 'Other British West Africa' 1·1
11–20	24	28·5	South Africa 8·5 U.S.A. 4·4 France 3·3 British Malaya 2·4 Belgium 1·8 U.S.S.R. 1·4 Poland 1·1 Misc. European countries 2·3
0–10	37	14·7	Germany 4·4 Italy 1·2 Brazil 1·1 Misc. Far Eastern countries 3·6 Misc. European countries 2·5 French, Portuguese, and Belgian Africa 1·0
	98	100·0	

---

[*] Taken from *Economic Journal*, March 1946, with slight revisions.

[†] Gold exports (or gold output if lower) are included in the exports of gold-producing countries, but not in their exports to the United Kingdom, since the sale of the gold did not depend on our market.

true, a fair number of unimportant trading nations, but they nevertheless took between them 43 per cent. of our pre-war exports. There were also four countries that, while depending on the United Kingdom market for more than one-fifth of their exports, depended even more on the market of the United States, and might be unwilling to yield to pressure by us for fear of offending that country. If these four (of which Canada is the most important) are added to our list of sixty-three, we get a group of sixty-seven 'unexploitable' countries which took roughly one-half of our exports before the war and supplied nearly one-half of our imports.

The conclusions so far seem to be (1) that while the number of countries to whom our market is of overwhelming importance is very small indeed, there is a fair number of important trading countries to whom our market is of very considerable importance; and (2) that the majority of the countries of the world, countries with which we did about half our trade before the war, would not be greatly susceptible to the use of our bargaining power as a large importer.

If, following my earlier nomenclature, we call this latter group of countries 'unexploitable' and the rest 'exploitable', we may draw up Table III, showing the geographical distribution of the two types of country.

TABLE III*

'Exploitable' countries	*Percentage of all U.K. exports in 1938 sent to 'exploitable' countries*	*'Unexploitable' countries*	*Percentage of all U.K. exports in 1938 sent to 'unexploitable' countries*
Scandinavia, Baltic States, and Holland	12½	Russia, and most European countries south of Baltic	18
Australia and New Zealand	12	Most of North and South America	12
India, Burma, and Ceylon	9	Union of South Africa	8½
Argentina and Uruguay	5	Most Far Eastern countries	6
Eire	4		
Miscellaneous	8	Miscellaneous, mainly African countries	5
	50½		49½

Notes: (1) 'Unexploitable' countries are those sending one-fifth or less of their exports to the United Kingdom in 1938, and other countries for which the United States was a larger market. 'Exploitable' countries include all others.

(2) Gold exports (or gold output if lower) are deemed to go to the United States.

* Taken from *Economic Journal*, March 1946.

Any attempt to force our exports by the use of our bargaining power as a large importer would be largely confined to the 'exploitable' countries. It may be that we should add one or two parts of the Colonial Empire, such as Malaya, which, although not greatly dependent on the United Kingdom market, might be controlled by us. But it is also true that the countries labelled ‘exploitable’ include the pre-war Baltic States, now part of Russia; other countries where our political power is uncertain; and Holland, which is forming a Customs Union with one of our industrial rivals. On balance, it is unlikely that the truly ‘exploitable’ countries took more than half our exports before the war.

To secure the necessary increase in our total exports by the use of our bargaining power would thus require a very large increase in exports to the ‘exploitable’ countries. This might, in principle, be achieved by an increase in their total imports of manufactures. But the exercise of our bargaining power would be likely to aim also at an increase in their imports of British manufactures at the expense of the manufactures of our rivals (*a*) in the United States, (*b*) in Europe.

(*a*) The United States did not, before the war, export very much to the ‘exploitable’ countries (some relevant figures will be found in Table IV), so that only a drastic discrimination by those countries against United States’ goods and in favour of ours would be of appreciable benefit to our exports. This would raise the danger, at least in the long run, of retaliation by the United States which would endanger a large part of our markets in North and South America, the Far East, possibly South Africa, and perhaps other countries dependent on financial assistance from the United States. These markets (as can be seen from Table III) took a good quarter of our exports.

TABLE IV*

	Imports (millions of dollars) in 1938 from					
	U.K.	*U.S.*	*Japan*	*Germany*	*Other industrial Europe*	*All main industrial countries*
Australia . . .	224	82	23	21	32	382
New Zealand . .	106	26	5	4	9	150
Eire. . . .	102	23	2	7	16	150
India, Burma, Ceylon	225	47	69	52	59	452
British Malaya† .	59	10	7	6	13	95
African Colonies .	86	17	9	11	18	141
Denmark . . .	123	29	—	87	59	298
Norway . . .	68	32	2	50	83	235
Portugal . . .	17	12	1	17	24	71
Argentina. . .	89	77	17	45	106	334
Total of above	1,099	355	135	300	419	2,308

* Source: *Network of World Trade.*

† British Malaya is included with other ‘exploitable’ countries because she is part of the Colonial Empire.

(*b*) We might in addition (or as an alternative) ask the ‘exploitable’ countries to buy from us rather than from our rivals in Europe. We might gain some of the pre-war German markets without provoking retaliation, but if we attempted to secure the pre-war export markets of the other European industrial countries (France, Belgium, Switzerland, Sweden, &c.) which in the aggregate were even more important than the German, we should seriously reduce their importing power. This might well impel them not only to reduce their purchases from the United Kingdom but also, perhaps, to enter into preferential arrangements with the other countries of Europe (including possibly some of the countries labelled 'exploitable'). We should then find our exports restricted, not only in eastern Europe, but also in central and western Europe.

If we used our bargaining power to gain exports in the 'exploitable' countries by forcing them to discriminate against the goods of our rivals, in both Europe and the United States, we should thus risk the loss of markets throughout the 'unexploitable' world, as shown on the right-hand side of Table III. Since many of the 'exploitable' countries were already taking a high proportion of their manufactures from the United Kingdom before the war (the 'exploitable' countries in the Empire took about one-half) the gains to be made are limited, while the possible indirect losses are large.

It would be very difficult to sell the exports necessary to pay for our essential imports if our trade were largely concentrated on a group of countries roughly corresponding to those labelled 'exploitable'. Even if we could sell the volume of exports that would, in a world undivided into economic groups or spheres of influence, pay for our essential imports, it is doubtful whether this would suffice if we relied heavily on our trade with the 'exploitable' countries. This group of countries supplied, it is true, over 80 per cent. of our meat, dairy produce, wool, tea, jute, and wood-pulp. This suggests, at first sight, that we could secure a large part of our essential imports by using our bargaining power with them. But they supplied only about one-third of all our other imports, and a very much smaller proportion of certain important commodities. A concentration of our trade on the countries with whom our bargaining power is greatest would, therefore, leave us deficient in many vital imported supplies, if our trade with the rest of the world was, as a result, greatly reduced.

This danger, which does not depend on the assumption of inconvertible currencies, may be illustrated by the following example (in which the figures are purely illustrative). The 'exploitable' countries took about half our exports before the war. If, by the use of our bargaining power, we induced them to take twice as great a volume as before, while, as a result of retaliation and other indirect consequences, our exports to the 'unexploitable' countries fell by one-half, our total exports would be 125 per cent. of the pre-war volume. Let us now suppose that this would enable us to buy 75 per cent. of the pre-war volume of imports, and that such a volume would just suffice to maintain a tolerable standard of living, *if we were free to buy throughout the world, and to apportion our import programme without regard to commercial agreements with particular countries*. But if we were not so free because, in order to push our exports in the 'exploitable' countries, we had agreed to take more from them than we should otherwise have done with a total import volume 75 per cent. of pre-war, then our imports would *not* suffice to maintain a tolerable standard of living. We might have relatively plentiful supplies of meat, dairy produce, wool, &c., but this would mean that we had not sufficient foreign exchange left over to buy even the minimum necessary supplies of many other commodities. This would be especially true if we had diverted our demand for these other commodities from the 'unexploitable' to the 'exploitable' countries, and if the latter could supply the diverted quantities only at a higher price.

3. Individual Countries' Exports to the United Kingdom as a Proportion of their National Income

The main criterion of our bargaining power used in section 2 was the proportion of a country's exports sent to the United Kingdom. This gives some indication of its dependence on our market for the balancing of its international accounts and for its power to buy imports. But it will usually give an exaggerated idea of how far its standard of living and level of employment depend on exports to the United Kingdom. If a country that sent 10 per cent. of its exports to us lost the whole of our market, then, even if it gained no compensating markets elsewhere, it is unlikely that its level of employment or standard of living would fall by as much as 10 per cent., because only a fraction of its resources are employed in production for export. Suppose this fraction were one-fifth. The fall in employment or living standards would then be likely to be nearer 2 than 10 per cent.

The fall in employment might, it is true, exceed 2 per cent. if the cumulative effect of a fall in export employment on employment in other industries were not offset. And, even if the displaced export workers were re-employed in production for the home market (and the old level of total employment maintained), the standard of living, in so far as it can be measured, might fall by more than 2 per cent. For the import economies that would be necessary might involve key raw materials or highly prized finished goods that could not be easily replaced by home production. In less unfavourable circumstances, however, if imports could be easily replaced by home production, the standard of living might equally well fall by less than 2 per cent. In general, the proportion of a country's *exports* sent to the United Kingdom is likely to exaggerate the dependence of its employment and standard of living on our market. The proportion of its total national *output* exported to the United Kingdom will, in some ways, give a truer picture.

Table V shows, in column 1, the exports of a number of countries in 1938 as an approximate percentage of their national incomes. This will overstate the percentage of the national effort going to export in so far as imports of raw materials and semi-processed goods are used in the production of exports; but it will understate the country's dependence on receipts from abroad by the amount of invisible exports. Many of the figures are not comparable with each other and some are subject to a wide margin of error, but the table probably gives a correct general impression of the relative importance of exports in the various countries' national outputs. The proportions shown in column 1 vary from 5 per cent. or less for the U.S.A., China, and almost certainly Russia, to one-third in New Zealand and Belgium and nearly one-half in Ceylon. Column 3 shows each country's exports to the United Kingdom as an approximate percentage of its national income. This proportion varies from 1 per cent. or less in a large number of countries to one-quarter in Ceylon and New Zealand. In the majority of countries it is 5 per cent. or less. The table includes countries with which we did about 70 per cent. of our trade in 1938. Making allowance for other countries it seems that about 60 per cent. of our exports before the war went to countries whose exports to us were 5 per cent. or less of their national incomes.

This reminds us again of the limited nature of our bargaining power with many important countries. It also suggests that, in exercising what power we have, we should often have to rely more on the importance of our market to a small section of producers than on its importance to a country as a whole.

Table V also shows that the proportion of their total exports sent to the United Kingdom by the various countries (column 2) usually indicates the relative dependence of their economies on the United Kingdom market, as shown by column 3. There are, however, important exceptions. Belgium and South Africa, for example, depended more on our market than did the United States, Poland, or Russia, although the last three countries sent a higher proportion of their exports to us.

4. The Proportion of World Exports of Individual Commodities Taken by the United Kingdom

In some ways the proportion of world exports of particular commodities taken by the United Kingdom is a better indication of our bargaining power than the proportion we take of each country's total exports. This will be the case:

TABLE V

	Total exports as approx. percentage of National Income, 1938 (c) (1)	*Exports to U.K. as approx. percentage of:*	
		Total exports, 1938 (2)	*National Income, 1938* (3)
Ceylon	48	54	26
New Zealand	31	83	26
Bolivia (*a*)	27	63	17
Denmark	26	56	15
Ireland	16	93	15
Finland	29	45	13
Canada	23	37	9
Australia	18	49	9
Dominican Republic (*a*)	16	40	7
Chile (*a*)	24	23	6
Norway	21	28	6
Netherlands	21	23	5
Argentina	16	33	5
Belgium	33	14	4
Sweden	16	24	4
South Africa	23	13	3
Brazil (*a*)	23	9	2
Czechoslovakia	18	9	2
Switzerland	16	11	2
Philippines	29	3	1
Netherlands East Indies	24	5	1
Greece	17	9	1
Japan (*b*)	16	5	1
Mexico	12	12	1
Hungary	10	7	1
France	9	12	1
Italy (*a*)	9	6	1
Poland (*a*)	7	19	1
U.S.A.	5	17	1
Bulgaria (*a*)	10	4	—
Germany (*b*)	7	6	—
China (*a*)	2	8	—
U.S.S.R.	..	20	(Under 1)

Notes: Gold exports (or output if lower) are included in the exports of gold-producing countries, but not in their exports to the United Kingdom.

Sources: Except where otherwise noted the national income figures are from the United Nations' *Monthly Bulletin of Statistics*, q.v. for definitions. The export figures are from *The Network of World Trade*, and the gold statistics from the sources stated in *Economic Journal*, March 1946, p. 36.

(*a*) National income figures derived from p. 243 of United Nations' *Economic Report*, 1945–7.

(*b*) National income figures from League of Nations' *World Economic Survey*, 1938/9, p. 84.

(*c*) The corresponding figure for the United Kingdom is 10 per cent.

(*a*) *When the proportion of a country's produce sent to the United Kingdom could very easily be modified.*

Thus we may take a large proportion of the world's exports of a commodity, but a small proportion of a particular country's exports because it is slightly more convenient for the country to send its produce elsewhere. For example, we took one-third of the world's exports of wheat and wheat-flour before the war, but only about one-sixth of the Argentine's wheat and flour exports. But the Argentine was greatly dependent on British purchases of wheat since these had a major influence on the world price. The proportion of its total exports sent to the United Kingdom by the Argentine tended, for this reason, to understate its dependence on the United Kingdom market; although there may, of course, have been offsetting tendencies.

On the other hand, the high proportion of Bolivia's total exports sent to the United Kingdom (over three-fifths in 1938) may have exaggerated her dependence on this country. Bolivia sent three-quarters of her tin ore (accounting for 70 per cent. of her total exports) to the United Kingdom for smelting. But the U.S.A. was a far larger consumer of tin than the United Kingdom and, although she did no smelting, her consumption had a great effect on the demand for, and price of, tin and so of tin ore.

(*b*) *When there are influential 'lobbies' of producers of particular products in overseas countries.*

Although a country sends a small proportion of her total exports to the United Kingdom, she may send us a high proportion of her exports of a particular product, the producers of which may be able to influence the commercial policy of their government.

In other ways the proportion of each country's total exports sent to the United Kingdom may often be the better measure of its dependence on our market.

(*a*) *When the world market for a product is imperfect, because of high transport costs, differences in quality, or for other reasons,* the proportion of a country's exports sent to the United Kingdom may be a better measure of its dependence on us than the proportion we take of world exports of the commodity. Suppose Denmark exported nothing but eggs. Then the fact that she sent nearly three-quarters of her pre-war exports of eggs to the United Kingdom would be a more accurate measure of her dependence on our market than the fact that we took only half the world's egg exports, if it were established that other markets for eggs were not so easily accessible to Danish exporters.

(*b*) *Even when a country sends a high proportion of its exports of certain commodities to the United Kingdom, it may send us a low proportion of its exports of many other things.* In this case the proportion of its total exports sent to the United Kingdom will be a better guide to its reliance on the United Kingdom market, at least for the balancing of its international accounts and for its power to import. Provided there are no powerful 'lobbies', this proportion will also be a better indication of our bargaining power with the country.

With these considerations in mind, we may examine the proportion of world exports of particular commodities taken by the United Kingdom before the war. This is shown in Table VI for fifty-one commodities. These accounted for two-thirds, by value, of Britain's imports before the war, but will probably account for a higher proportion in the future since it is likely that our imports of manufactures will be severely restricted.

It will be seen that the United Kingdom took a very high proportion of world exports of a number of important commodities, especially animal products and tea; and, as we saw in section 2, we took a high proportion of the total exports of countries that specialized in these commodities, notably Australia, New Zealand, Argentina, Eire, Denmark, India, and Ceylon.

But there were also many commodities of which the United Kingdom took only a small proportion of world exports. She took less than a fifth of world exports of rice, oats, potatoes, wine, coffee, cocoa beans, bananas, dates, oilseeds, vegetable oils, pepper, hops, kidney beans, rubber, silk, flax, copper, tin, bauxite, iron ore, manganese ore, crude petroleum, gas and fuel oil, cattle hides, nitrates, and potash fertilizers. For the majority of these commodities the United States was a bigger market. If we add to the list sugar and wood-pulp, for which, although the United Kingdom share was one-fifth or more, the United States was a still larger market, we get 28 commodities out of 51 for which the pull of the United Kingdom market cannot be regarded as the predominant factor.

By analogy with section 2 we might perhaps say that the producers of over half the commodities on our import list are 'unexploitable'. This would, of course, be misleading in so far as the twenty-eight items cost us only £124 million out of a total import bill of £571 million for the full list of fifty-one items. Over three-quarters, by value, of the imports in Table VI were thus of commodities for which the United Kingdom market was of great importance. Our bargaining power appears, at first sight, to be much greater when we take each commodity separately than it was when we made the analysis by countries. If we were to use our bargaining power with the producers of the 'exploitable' commodities to make them take payment for our purchases in British goods, could we not secure the great bulk of our essential imports?

TABLE VI

	Pre-war share in world imports		*U.K. retained imports, 1938* (*£ millions*) (3)	S = *U.K. imports less than 20% of world total.* U = *U.K. imports less than U.S. imports* (4)	
	U.K. % (1)	*U.S.* % (2)			
FOOD: Meat (*a*)	80	3	90	..	.
Butter	78	1	50	..	..
Eggs	52	—	15	..	..
Cheese	52	10	10	..	..
Wheat (incl. flour)	32	3	43	..	..
Rice	1	1	1	S	..
Maize	27	16 (*b*)	18	..	..
Barley	34	8	7	..	..
Oats	8	—	1	S	..
Potatoes	16	2	2	S	..
Sugar	20	42 (*c*)	19	..	U
Wine	5	1	5	S	.
Coffee	1	47	1	S	U
Tea	51	10	26	..	..
Cocóa beans	14	41	4	S	U
Citrus fruit	39	1	10	..	..
Bananas	12	55	5	S	U
Apples	45	—	6	..	..
Dates	9	11	—	S	U
Raisins	44	1	3	..	..
Grapes	21	2	2	..	..
Onions	47	—	2	..	..
6 Oilseeds (fat content) (*a*)	16	10	14	S	..
8 Vegetable oils (*a*)	14	25	4	S	U
Pepper	7	29	—	S	U
Hops	11	20	—	S	U
Dried kidney beans	7	5	—	S	..
MATERIALS: Timber	34	6	43	..	..
Wood-pulp	28	34	14	..	U
Rubber	10	43	9	S	U
Wool	30	13	28	..	..
Cotton	22	2	28	..	..
Silk	5	64	2	S	U
Jute	21	13	3	..	..
Hemp	28	12	2	..	..
Flax	16	2	4	S	..
Tobacco	22	6	22	..	..
Copper (unworked) (*a*)	18	11	11	S	..
Lead (unworked) (*a*)	47	1	6	..	..
Tin (unworked) (*a*)	6	40	1	S	U
Tin ore (*a*)	40	—	6	..	..
Zinc (unworked) (*a*)	36	1	2	..	..
Bauxite (*a*)	11	20	—	S	U
Iron ore (*a*)	11	5	7	S	..
Manganese ore (*a*)	9	23	1	S	U
Crude petroleum (*a*)	4	7	5	S	U
Petrol (*a*)	31	—	25	..	..
Gas and fuel oil (*a*)	18	15	9	S	..
Cattle hides	14	10	4	S	..
Nitrates	2	30	—	S	U
Potash fertilizers (*a*)	9	14	1	S	U
Total of above	..	..	571		
Other Commodities	..	..	287		
ALL COMMODITIES	17	12	858		

Sources: *International Yearbook of Agricultural Statistics* (International Institute of Agriculture); *International Trade in Certain Raw Materials and Foodstuffs* and *Network of World Trade* (League of Nations); *Annual Statement of the Trade of the U.K.*

Note: The figures in columns 1 and 2 refer to 1937, except those for commodities marked (*a*) which refer to 1938. For commodities marked (*a*) the figures in column 1 include imports into Eire, Iceland, the Faroe Islands, and Spitzbergen.

(*b*) The large import of maize into the U.S. in 1937 was abnormal.

(*c*) Including imports from Puerto Rico, Hawaii, and U.S. Virgin Islands.

But although the twenty-eight items were less than one-quarter, by value, of *Britain*'s imports, they formed a much larger proportion—roughly one-half—of *world* trade in the fifty-one commodities. The producers of the twenty-eight commodities were therefore important *purchasers* of manufactured goods in the world market. If we add the exporters of coal (of which we imported none) and of manufactures (of which we imported only 8 per cent. of the world total) it will be seen that a large part of the world's producers (and so a large part of the world's purchasers) were not greatly dependent on the United Kingdom market. But we depend greatly on their purchases of our manufactured goods. And since countries in which these producers are the most influential would not be very susceptible to pressure by us, the commodity analysis is quite consistent with the country analysis, which showed only one-half of our exports going to countries likely to be 'exploitable'.

5. The Small Number of Suppliers of Individual Commodities Imported by the United Kingdom

Even where we take a substantial proportion of world exports of a particular commodity our bargaining power may not be very great if the number of suppliers is small. This is in large measure the case. Table VII shows the percentage of United Kingdom imports of various commodities drawn from our main suppliers, defined as those countries that, between them, sent us 75 per cent. or more of our imports of each commodity. The table includes most of the commodities in Table VI of which we took more than one-fifth of world exports, and a number of others as well. It shows that at least three-quarters of our supplies of nearly every commodity came from a very small number of countries. Our most important single supplier usually sent a very high proportion of our total imports; always more than one-quarter and very often more than one-half.

This means, it is true, that only a small number of negotiations would be needed to cover the greater part of our requirements of each commodity. But it also means, first, that it would be easy for the main suppliers to co-operate in order to resist our demands; indeed, any serious attempt to use our bargaining power might lead to agreements that would not otherwise have been made, and to a higher price than would have had to be paid if we had refrained from such tactics. The small number of suppliers means, in the second place, that our bargaining power, with the more important at least, will be smaller than is suggested by the proportion of their exports sent to this country. It is most unlikely, for example, that we could dispense with the whole of our tea imports from India, the whole of our beef imports from the Argentine, or the whole of our tobacco

TABLE VII

Proportion of United Kingdom Imports Obtained from Main Suppliers in 1938

Commodities of which ONE *country supplied 75 per cent. or more of U.K. imports*

		Per cent.
Nickel ore . .	Canada	100
Nickel (unwrought) .	Canada	90
Jute . . .	India	100
Molybdenum ore .	U.S.A.	97
Tobacco . .	U.S.A.	76

Commodities of which TWO *countries supplied 75 per cent. or more of U.K. imports*

		Per cent.		Per cent.
Flax . .	Belgium	48	U.S.S.R.*	42
Tea . .	India	55	Ceylon	32
Rubber . .	Malaya	68	Dutch E. Indies	14
Mutton and lamb	N. Zealand	52	Australia	27
Beef and veal .	Argentina	60	Australia	18
Cheese . .	N. Zealand	56	Canada	23
Tin ore . .	Bolivia	55	Nigeria	20
Wood-pulp .	Finland	40	Sweden	35
Hemp . .	Philippines	46	Brit. E. Africa	32
Zinc (unworked) .	Canada	54	Belgium	22

Commodities of which THREE *countries supplied 75 per cent. or more of U.K. imports*

		Per cent.		Per cent.		Per cent.
Bacon .	Denmark	49	Canada	19	Holland	7
Wheat .	Australia	31	Canada	28	U.S.A.	16
Wool . .	Australia	41	N. Zealand	23	S. Africa	12
Lead . .	Australia	46	Canada	26	Burma	15
Maize .	U.S.A.	40	Argentina	31	Canada	6
Cotton .	U.S.A.	34	Egypt†	34	India	12
Motor spirit	Dutch W. Ind.	41	U.S.A.	22	Iran	22

Commodities of which FOUR *or* MORE *countries supplied 75 per cent. or more of U.K. imports*

		Per cent.		Per cent.		Per cent.		Per cent.
Butter .	N. Zealand	27	Denmark	25	Australia	19	Holland	7
Barley .	Canada	27	U.S.S.R.	19	Iraq	16	U.S.A.	14
Timber	U.S.S.R.*	37	Canada	17	Sweden	14	U.S.A.	9
Oranges	Palestine	37	S. Africa	15	Brazil	15	U.S.A.	14
Eggs .	Denmark	34	Holland	21	Poland	9	Eire Rumania	8 5

* Including pre-war Baltic States.

† Including Sudan.

Source: *Annual Statement of the Trade of the U.K.*

imports from the United States. Our bargaining power with a country, or with a section of producers, is determined not so much by the proportion of their exports they have hitherto sent us, but by the proportion we might dispense with.

Even when we could dispense with the bulk of our imports from country A, by taking them from country B instead, this would be little hardship to country A if country B's exports could not be greatly expanded; A would merely divert her exports to the markets now vacated by B. The loss of exports with which we could threaten A would then be limited roughly to the possible increase in B's exports. But if A's exports form a large proportion of the world total this possible loss will not normally be of great importance to her, since it is usually difficult to make a rapid expansion in primary production.

Even, then, when we take a high proportion of the world's exports of a commodity, our bargaining power will be circumscribed if the available export surpluses are in the hands of a small number of countries; for this will reduce our power to play off one primary producer against another. In the limiting cases of, say, jute or Manila hemp, of which India and the Philippines, respectively, have virtual world monopolies, we can reduce our purchases only by reducing total consumption. There is no alternative source of supply unless other countries begin to produce.

We can, of course, substitute, up to a point, one commodity for another —other textiles for jute and Manila hemp; beef for mutton; even wheat and maize for bacon and eggs. But when we analyse our imports in broad groups, we still find a marked concentration of supplies in the hands of a small number of countries. Thus, for example, six countries (Australia, New Zealand, Denmark, Argentina, Canada, and the U.S.A.) supplied between them three-quarters of our imports of cereals, meat, and dairy produce of all kinds.

It might be argued that the high proportion of *United Kingdom imports* of particular commodities obtained from a small number of countries exaggerates the concentration of *world exports,* since we may take our imports largely from countries with which we have close trading relations. But this is unlikely to be an important source of error where we take a high proportion of world imports; and in fact world exports of most primary commodities are highly concentrated. This is illustrated in Table VIII. The last column of figures shows that, nearly as often as not, the four main exporters supply four-fifths or more of the world total. The first column shows that the largest single exporter always supplies one-quarter of the total and often two-fifths or more. This is true even though some of the categories are wide. The commodities included under cotton, wood-pulp, wool, tobacco, and hemp, for example, are by no means perfect

TABLE VIII

	Percentage of 1937 World Exports supplied by four main exporters of each commodity								*Total of four countries*
Jute	India	97	Nepal	1	Belgium	1	China	1	100
Raw silk	Japan	83	China†	11	Italy	5½	France	½	100
Mutton	New Zealand	52	Australia	27	Argentina	15	Uruguay	2	96
Rubber	Malaya	40*	Netherlands E. Ind.	40	Ceylon	6	Indo-China	4	90
Hemp	Philippines	49	Italy	15	India	14	Yugoslavia	11	89
Cotton	U.S.A.	45	India	21	Egypt§	15	Brazil	7	88
Flax	France	39	U.S.S.R.‡	19	Netherlands	17	Belgium	12	87
Maize	Argentina	70	S. Africa	6	Yugoslavia	6	Indo-China	4	86
Wood-pulp	Sweden	40	Finland	23	Canada	12	Norway	10	85
Rice	Burma	33	Indo-China	23	Thailand	16	India	11	83
Tea	India	32	Ceylon	20	Netherlands E. Ind.	16	Brazil	14	82
Beef	Argentina	50	Australia	13	Uruguay	10	Brazil	6	79
Cocoa beans	Gold Coast	41	Brazil	16	Nigeria	14	French West Africa	8	79
Coffee	Brazil	49	Colombia	17	Netherlands E. Ind.	7	El Salvador	5	78
Motor spirit	U.S.A.	33	Curaçao	18	Netherlands E. Ind.	13	Iran	12	76
Butter	Denmark	26	New Zealand	25	Australia	14	Netherlands	9	74
Cheese	New Zealand	29	Netherlands	22	Canada	14	Italy	8	73
Wool	Australia	36	Argentina	12	New Zealand	11	S. Africa	11	70
Wheat	Argentina	28	Canada	19	Australia	15	Rumania	7	69
Pig meat	Denmark	34	Canada	18	U.S.A.	7	Poland	7	66
Tobacco	U.S.A.	38	Netherlands E. Ind.	10	Greece	8	Turkey	8	64
Eggs	Denmark	26	Netherlands	19	Poland	7	China	6	58
Sugar	Cuba	29	Netherlands E. Ind.	13	Philippines	10	Australia	5	57

* Exports of home-produced rubber only. † Including Manchuria. ‡ Including pre-war Baltic States. § Including Sudan.

Sources: *International Trade in Certain Raw Materials and Foodstuffs* and *International Year Book of Agricultural Statistics.*

substitutes. Had a finer classification been used the concentration would have been even greater.

It may be argued that, even where there is a small number of actual exporters of a commodity, there may also be potential exporters, and that this would increase our bargaining power with the actual exporters. But if the switching of our purchases to countries not yet producing, or not yet exporting, involved a large increase in the price of our imports, our bargaining power would not in fact be so great. If, on the other hand, it is held to be easy for primary producers to expand production at a cost not much greater than the old world price, then it will in general be easy for existing producers of one commodity to switch to the production of another. This would also limit the effectiveness of a threat by us to reduce our purchases.

6. Proportion of the World's, and of Single Countries', Production of Individual Commodities Exported to the United Kingdom

When we take a high proportion of world *exports* of a commodity, it is salutary to remember the much smaller proportion of world *production* dependent on our market. This is illustrated in Table IX.

TABLE IX

	United Kingdom imports in 1937 as percentage of	
	World exports (1)	*World production* (2)
Butter	78	11
Beef and veal	61	4
Cheese	52	7
Barley	34	2
Wheat	32	4
Maize	27	3
Tobacco	22	4
Cotton	22	9
Sugar	20	8

Sources: *International Year Book of Agricultural Statistics*; League of Nations' *Statistical Year Book* and *International Trade in Certain Raw Materials and Foodstuffs*; *Commodity Year Book*.

Note: Some of the figures in column 2 may be too high because of the incompleteness of world production figures.

For the major exporting countries, it is true, the proportion of output exported to the United Kingdom is often much larger than the world proportion shown in Table IX. This is illustrated in column 1 of Table X. But column 3 of the same table reminds us that even the major exporting

TABLE X

		Percentage of pre-war production			
		Exported to U.K. (1)	*Exported to other countries* (2)	*Used at home** (3)	*Total* (4)
Beef and veal	Argentina	20	5	75	100
	Australia	20	2	78	100
Mutton and lamb	New Zealand	73	—	27	100
	Australia	27	1	72	100
	Argentina	28	1	71	100
Pig meat	Denmark	58	4	38	100
	Canada	29	3	68	100
Butter	Denmark	62	21	17	100
	New Zealand	82	2	16	100
	Australia	42	4	54	100
Cheese	New Zealand	90	—	10	100
	Netherlands	9	41	50	100
	Canada	61	6	33	100
Wheat†	Argentina	5	29	66	100
	Canada	44	23	33	100
	Australia	36	30	34	100
Tobacco	U.S.A.	17	16	67	100
Jute	India	12	38	50	100
Cotton	U.S.A.	8	31	61	100
	India and Burma	9	50	41	100
	Egypt	30	57	13	100

* Including changes in stocks, quantities used for seed, &c.

† Exports include wheat equivalent of flour exports.

Notes: (1) The figures refer either to a single recent pre-war year, or, where it appears that this would be misleading, to an average of such years.

(2) The countries shown are those that supplied more than 10 per cent. of world exports of each commodity in 1937, as shown in Table VIII.

(3) The figures for meat may exaggerate the proportion of output used at home in so far as a higher proportion of the tonnage exported than of the tonnage produced consists of edible meat. The production figures used refer to dressed carcases, i.e. excluding head, extremities, and offal.

Sources: As for Table IX, together with certain national statistical sources.

countries, the only ones shown, often consume a high proportion of their output themselves. Thus the United States consumed two-thirds of her tobacco production. Even Australia and the Argentine, two of our major suppliers of meat, seem to have consumed about three-quarters of their own output, although the figures may exaggerate the true proportion.[1]

---

[1] See note 3 to Table X.

This reminds us of the limitation of our bargaining power when aimed at producing interests in overseas countries. Even a cut of, say, one-third in our imports of meat from Australia or the Argentine would affect the total meat production in these countries by only about one-tenth. This could be fairly rapidly offset by increased domestic consumption, by the development of other branches of agriculture, or by more rapid industrialization.

7. Special Characteristics of the Post-war Position

Unless there is a large expansion in world trade, British imports, especially of foodstuffs, are likely to remain well below the pre-war level, and well below the level we should like to take if our importing power were greater.

In some ways this may increase our bargaining strength. In the first place we shall be more willing to substitute one commodity for another. Secondly, we shall be more ready to offer to increase our imports of food from a country that gives us adequate concessions, since this will be less likely to harm the interests of British farmers. Thirdly, the low exports of our suppliers, resulting from our reduced importing power, may make them more anxious for a large share in the British market, and so more willing to give favourable treatment to our exports.

But in other ways our bargaining power may be reduced. In the first place, the more our main suppliers are depressed, the more likely they are to make agreements with each other, and this would reduce our bargaining power.

Secondly, it is by no means certain that, even if our imports are greatly curtailed, the primary producers will be in desperate straits. We must remember that our loss, through reductions in net invisible income and possibly through adverse terms of trade, will be their gain. Even, then, if we are importing from them a considerably lower volume of goods than before the war, they may be importing just as great, or possibly a greater volume from us.

Thirdly, our main suppliers have continued to industrialize during the war. This will make them less dependent on imported manufactures, and also help to provide employment for those displaced from primary production as a result of a fall in exports to Britain.

Finally, any fall in exports to Britain may well be offset by additional sales elsewhere. In the early post-war years, at least, there has been no surplus food in the world. The main reason has been the fall in production which it is hoped will be temporary. But the growth of population, together with full employment policies, are likely to keep the demand for primary products at a higher level than before the war, both within the countries exporting these products and in other countries. In particular, the level of industrial production in the United States is almost certain to be much higher than before the war, and this will greatly affect the world demand for primary products.

For all these reasons, our suppliers may not all be so desperately anxious to sell additional exports, or to get a better price, that they are willing to give very favourable terms to our exports.

We, on the other hand, shall be extremely anxious about our balance of payments. In bargaining with, say, the Argentine, we should be susceptible to a threat by them to buy more from the United States and less from us, even if this would mean quite a small fall in our exports, and so in our imports. A given proportionate fall in our imports will be a much more serious matter when they have already been cut to the bone than it would have been before the war. Our suppliers, on the other hand, may be more easily able to accept a cut in their exports.

We should never forget that, in the last resort, the bargaining power of a manufacturing nation like Britain, with very little agricultural land per head of the population, and with virtually no other natural resources save coal and iron, may well be weaker than that of primary producing nations, especially advanced countries like Australia, New Zealand, the Argentine, Canada, or Denmark. For whereas the latter can, if necessary, expand their industrial production without much difficulty and so dispense with our exports, we are far less able to expand our production of food and raw materials.

8. Conclusion and Summary

The general conclusion of this note is that our bargaining power as a large importer is not so great as is sometimes thought. There are many countries—important markets for our exports—that do not depend greatly on the United Kingdom market for the sale of their exports, or that depend even more on the market of the United States; and nearly all countries depend even less on our market for the maintenance of their standard of living and level of employment. Thus many countries would not be susceptible to pressure by us.

It is true that we take a high proportion of world exports of many important commodities. We might hope that their producers would bring pressure to bear on their governments to give concessions to Britain. But there are many other important commodities of which Britain takes only a small proportion of world exports. Although these are relatively unimportant in our import pattern, their producers form a large part of the world's purchasers of manufactured goods—our main export. Countries in which these producers are the most influential would not be very susceptible to pressure by us.

Even where we take a high proportion of world exports of a commodity, our bargaining power with its producers is not likely to be so great as appears at first sight, for two reasons.

First, a large part of the production is often consumed within the producing country. Secondly, we could often not wholly dispense with even the relatively small part exported to Britain, so dependent are we on a few countries for the bulk of our imports of most commodities.

Thus only a limited number of countries would be susceptible to our bargaining power as a large importer, and in many of these we have already exploited fairly fully the market for our manufactured goods. Our bargaining power, too, might in some instances prove short-lived. Nor is it likely to have increased on balance since 1938.

It does not follow, of course, that we should in no circumstances use what bargaining power we have. The general conclusion is merely that we should not exaggerate its importance when framing our commercial policy.

WADHAM COLLEGE,
OXFORD.

PROGRAMMES AND ALLOCATIONS IN THE PLANNED ECONOMY

By T. WILSON

I

In his lectures on *Central Planning and Control in War and Peace* Sir Oliver Franks offered the following definition of planning: 'Peace', he said, 'is very different from war and many of the war-time functions and forms of organization will find no application in peace. But the essence of planning and control by the State in relation to productive industry and commercial activity must be the same in war and peace. I think the essential elements are plans consisting of decisions of policy quantitatively expressed in the form of programmes and such measures as in particular circumstances may be necessary to ensure the performance of these programmes.' He goes on to add that: 'These elements were common to all the planning and controlling organizations of the war; they are discoverable in Government action at the present time. Without them I do not see how any Government could act as a planning and controlling authority.'

The definition is a neat one; indeed it is too neat for the untidy concept it is meant to describe. The term 'planning', like its shadowy twin 'co-ordination', is far too intangible to be captured in the philosopher's box. Everyone is agreed upon the need for planning, but there is far too much variety among the planners for any precise definition of these views. In particular a large number of planners, who differ appreciably among themselves, would be united in their rejection of Sir Oliver Franks's definition. Many of them believe in a great deal of State intervention, but they would give greater freedom to the consumer and would rely in the main on the pricing system to determine the allocation of resources. Sir Oliver has in fact described, not planning in general, as he seems to imply, but 'physical' or 'quantitative' planning.[1] The distribution of resources is to be largely settled by the State and enforced by means of controls. Prices and money are not, presumably, to be driven off the stage, but their role is rarely described and is presumably obscure.

The present article will be concerned with the technique of physical planning and it will follow more distinguished precedent by beginning with the war. No one will deny that when a very large proportion of the nation's resources is required by the State for its own use, liberal methods are out of place and more direct control of the composition of output is required. But there is good reason to criticize the tendency to assume that the preparation of output programmes is a fairly straightforward business, and still more reason to point to the much-neglected problem of estimating the resources required to implement these programmes. The belief that these tasks can be readily accomplished is ill founded and it is, therefore, all the more regrettable that Sir Oliver Franks says so little about these matters, especially the second, although they are crucial to his argument. Even Sir Hubert Henderson, who is rarely over-sanguine, passed over the difficulties a little too lightly in his Rede lecture on *The Uses and Abuses of Economic Planning*. If the agility of the 'invisible hand' has been called in question time and again, the dream of scientific planning deserves no better treatment.

---

[1] Both adjectives are unsatisfactory. 'Quantitative' implies that liberal planners, by contrast, pay no regard to figures. 'Physical' leads to a confusion with town and country planning. 'Authoritarian' would be the more accurate term, but since it would be resented —however unreasonably—I shall not use it, but will make 'physical' serve my purpose. It implies that decisions relate in the main to tons of materials and numbers of heads rather than to monetary quantities or prices.

The article will therefore begin with a brief reference to the manner in which munitions programmes were prepared during the war. This will be followed by a somewhat lengthy and detailed analysis of the various ways in which requirements were then calculated. Detail is essential at this point if the discussion is to consist of more than a string of bald assertions, but I am unhappily aware that the result is a few pages of quite uncompromising dullness. An attempt is then made to indicate the manner in which the government machine finally produced a list of the allocations granted to the various claimants. After this rough sketch I shall try to draw some conclusions with regard to physical planning under conditions of peace.

II

Little need be said in the present article about the initial step of preparing the production programmes, but it may be noted in passing that it was a far from simple task. The flow of components, produced under varying conditions in a large number of factories, had to be synchronized and great knowledge and skill were required to do so even with regard to a single type of tank or aircraft. If each article had been produced by one integrated firm the task could have been left to the management, but it was a different matter when a large number of firms were involved in each case, some engaged in the final assembly, some providing the radio valves, some the brakes, some the nuts and bolts, and so on. Bottlenecks of one kind or another were always causing a certain amount of trouble, although it is true to say that the synchronization of the supply not only of components but also of the various items making up complete sets of unit equipment was very much better done in the 1939–45 War than in the 1914–18 War. Partly because of the rearmament programme which had begun before the outbreak of war in 1939, a good standard was reached even in the early days of the war.

The simple method of getting synchronization is to keep all the supply bins filled, but this is an expensive process as it involves carrying large margins. When the war effort became stretched, it was too costly to carry store margins in excess of absolute need; and there tended accordingly to be more precise planning of the build-up of components and ancillary stores. Given the inherent difficulties, it was fairly well done —the proof being the relatively small excesses at the end. But the performance of this complicated feat involved much weary labour on the part of large staffs inside and outside the government service, and thus diverted much highly skilled labour from other tasks.

In the preparation of a programme the department concerned worked in close touch with 'the industry', and without such co-operation the job would have been hopeless from the outset. The industry, it should be noted in passing, meant the principal companies, for it was altogether out of the question to consult all the thousands of firms concerned. For this reason physical planning is bound to encourage the growth of monopoly, as was clearly recognized in 1919 by the Committee on Trusts.[1] The continued use of such methods is gradually transforming the United Kingdom into a corporative state, and it is no consolation that important groups of business men—who profess to dislike planning in theory—are by no means averse to such corporative developments so far as their own industries are concerned.

As Sir Oliver Franks has stressed, these programmes were not merely statements of desire; they were 'acts of will'. It was therefore necessary that they should be 'realistic' and, in fact, they became increasingly so as the years passed. To be realistic it had to be possible to produce them with existing capacity together with any additions on which it was reasonably safe to count. Some broad assumptions had also to be made about the supply of labour and raw materials, and although these were frequently found to be over-optimistic it was at least possible to ensure that they were not altogether absurd. Thus the programmes could be regarded as reasonable drafts, but they were no more than this until all the demands for the nation's scarce resources had been put together and a balance struck.

It is with this second stage of preparing a balance-sheet that we shall be primarily concerned. It need scarcely be said that the topic is far too large to be discussed comprehensively in an article and far too complex, in any case, to be handled by the present author. It may, however, be possible to indicate some of the main points by confining attention to one factor of production—man-power—and virtually ignoring all the rest. My comments will also refer to the last three years of the war.

---

[1] Cd. 9236, para. 4.

It is right, I think, that the choice should fall on man-power rather than raw materials, factory space, or machine tools. The man-power budget was always settled at the highest level, whereas the internal allocations of raw materials were left in the main to more humble committees. As for factory space and capital equipment, these rarely presented difficulties in the latter part of the war. It is true to say that the man-power decisions determined the size and the broad structure of the war effort and most of the other decisions, e.g. with regard to raw materials, were brought into harmony—mistakes apart—at a later stage. Labour was the most important of the internal, limiting factors, and it should also be observed that, although the supply of labour was far from homogeneous, it was sufficiently so to make possible the preparation of a budget which related to the whole economy. The man-power budget, moreover, has a final, if less serious, claim to be chosen because it has acquired such an odour of sanctity among the devotees of physical planning.

In what follows we shall therefore assume that the programmes have been prepared and ask how the interested departments calculated the man-power required for their production.

Two principal methods of calculation were employed, and these will be discussed in turn.

1. The first method was to ask the firms how much they required. Every quarter the Ministry of Labour put this question to a very large sample of producers and supplemented the results with those obtained from a smaller sample at more frequent intervals. In the case of the larger firms which were working on direct contract for the Supply Departments it was reasonable to expect a fairly satisfactory answer. These firms knew their programmes and had, indeed, co-operated with the civil servants in their preparation. They were accustomed to thinking for a long time ahead and had techniques for calculating requirements. Even in their case, however, there were obvious difficulties. Their future efficiency—which depended so much upon the quality of the new labour supplied, the possible loss of skilled men to the forces, the number of modifications demanded by the Service Departments, and so on—was to some extent a matter of guess-work. Furthermore, they were under a strong temptation to exaggerate their needs in anticipation of future cuts. It is true that firms with fixed-price contracts would not wish to be burdened with excessive man-power, but the danger of a permanent surplus probably seemed remote. In some areas, such as the Midlands, labour was so scarce and the flow of new recruits had been so inadequate for such a long time that a somewhat different problem arose. Not

unnaturally it was difficult for the managers in such areas to take seriously these academic questions about their requirements. They were inclined to suspect that any figure they submitted would have little effect on the numbers they obtained. Their day-to-day contacts with the Ministry of Labour and the Supply Department would ensure them a reasonable share of any man-power available and there was little point in wasting precious time on theoretical calculations.

There were serious difficulties but, in the case of the main contractors, the Ministry of Labour and the Supply Departments were sufficiently familiar with conditions in the firms to make some adjustment. It was a very different matter with regard to the host of sub-contractors, who numbered well over 20,000 in the case of aircraft alone. These firms knew little about the programmes; they merely knew what orders had been placed by the main contractors. Many of them were not in any case accustomed to making plans for a year or eighteen months ahead. In ordinary circumstances they would have employed a few extra workers at short notice and sacked them later if they proved to be redundant. In theory they might have been made to prepare long-term plans, but where was the technical staff necessary for such reforms to be found? In any case how could they be told about future programmes without incurring a serious security risk? Furthermore, many of them were working indirectly for more than one Supply Department and for the civilian market as well. They could not divide their requirements and were not, in fact, asked to do so. (The figure for global requirements, when finally tabulated, was split up by the Ministry of Labour and the interested Supply Departments according to an agreed but admittedly imprecise formula.) Labour scarcity and vacancies long unfilled often made these firms even more inclined than the main contractors to treat the questionnaires as academic. Their contact with regional offices was often very slight, they were less likely to see the connexion between the forms and the future supply of labour, and the departments had far less knowledge of their affairs with which to check the figures submitted.

If the number of firms was a major cause of the difficulties encountered, it also warranted the assumption that errors would tend to some extent to cancel out. Some firms, which took the forms seriously, overstated their needs in anticipation of cuts. Others put down figures more or less at random which were sometimes much too low, as detailed sample inquiries revealed.

In the end, it was felt that the figure for requirements over the next three months was likely to be of some value, but that estimates for the more distant future should be treated with great reserve. It was this time factor which proved to be a most serious limitation on the use of the method. A month had usually elapsed before the figures had been tabulated, submitted to certain tests, and analysed. The whole of the three months might well have passed before the Cabinet had reached a decision, especially if the man-power position was so difficult as to call for a revision of the programmes for the Services and the Supply Departments. Nevertheless the estimates provided a valuable check on conclusions reached by other methods, and by means of some manipulation and guess-work it was possible to project them forward beyond the three months. These adjustments were ingenious and a rough-and-ready answer could be reached. But the whole procedure was singularly lacking in that academic precision which 'planning' is supposed to offer in contrast to the 'chaos of the pricing system'.

2. The second method was to estimate productivity in the past and then calculate future requirements with the aid of the ratio thus determined. The various steps in the calculation will be discussed in turn, and aircraft production will be taken as an example.

(*a*) The first task was to put a figure on the actual labour force, and although this may appear quite straightforward, it was complicated in practice by the fact that so many firms were engaged on orders for other Supply Departments and for the civilian market. The Ministry of Labour issued forms each quarter to most of the firms, on which were to be recorded particulars about the labour employed and its distribution between different uses. These figures for distribution were usually based on the value of the work performed and therefore reflected some irrelevant features such as differences in contracting methods. Nevertheless the sample was sufficiently large to warrant the assumption that the errors would cancel out, at least for the country as a whole and for the larger regions. It was still necessary to gross up the sample figures, and this entailed some additional complications, but the final estimate for the man-power employed on production for each of the Supply Departments was probably near enough the mark. At any rate there was little reason to expect that the errors would differ from one period to another, and small errors, so long as they were consistent, were likely to be of negligible importance. In passing, however, it must be stressed that the making of these calculations was a big job. Forms were sent out to thousands of firms and by the time they were returned, tabulated, and analysed, the better part of a month had gone. These figures were available only once a quarter, and although it was possible to interpolate by means of much smaller samples for intervening dates, this involved additional labour and inconvenience to producers and civil servants alike and the results had to be treated with greater caution.

(*b*) The next step was much more tricky. There were far too many firms and there was far too much sub-contracting for the Ministry to calculate the total numbers employed on the production of each type of plane. The only possible procedure was to relate the figures for total man-power to the figures for total output, but output was far from homogeneous and a common denominator was therefore required. One method was to multiply the production of each type of aircraft by the structural weight of the airframe, but unfortunately it suffered from the defect that differences in structural weight were not directly correlated with differences in man-power. The correlation was imperfect even with regard to the airframe alone, and there was still less reason to suppose that this method would yield an accurate estimate of overall productivity, including the production of engines, propellers, under-carriages, and so on. To some extent it was possible to correct for these errors—an index of horse-power was also employed—and in this way an approximate answer was obtained.

An alternative method was to use as weights the estimates of man-hours required for the production of each plane. These figures were obtained from the Principal Technical Costs Officer and were prepared in the first instance for use in pre-costed, fixed-price contracts. On the face of it this common denominator appeared more satisfactory, but here too there were snags. The man-hours were calculated on the assumption of flow production and there were always a good many planes in the earlier stages of development. Even in the case of the older types, production was hampered by continuous modifications, and there was no reason to suppose that all planes were equally affected. Furthermore, the technical estimates implied 'reasonable' efficiency on the part of both management and workers, and here too there were marked differences as between one firm and another. For these reasons, the weights were clearly wrong, but, *faute de mieux*, it was appropriate to use them and to compare the results with those obtained by other methods.

(*c*) Fresh difficulties were encountered in attempting to relate the figures for man-power to those for output. It was clearly necessary to introduce some time-lag, but its accurate calculation would have required a good deal more knowledge about both the length and shape of the production periods than was in fact available. It was possible only to make a rough-and-ready guess for a few main types of plane and then make a further guess at the weighted average. This was, of course, a possible source of serious error. With both actual and planned production changing rapidly, variations in the amount of work-in-progress were of considerable moment. What may have been an appropriate time-lag at one stage in the war was not likely to be appropriate throughout, but elaborate adjustments were out of the question.

(*d*) Finally, when productivity in the recent past had been calculated, an estimate could be made of the labour force required at various dates to produce the programme, provided efficiency did not change. But efficiency was, of course, changing all the time. In the early years it was extraordinarily difficult to know what allowance to make for the entry of new firms with no experience of aircraft production or for the great dilution of the labour force. Nothing more than crude guess-work was possible, and many of the other defects in the calculations made were rightly regarded as trivial by comparison. As the years passed, the uncertainty was substantially reduced, although it still remained far too great for anyone to make confident predictions. On the whole productivity increased as the new labour and the new managements became more experienced, and teething troubles were left behind in the production of an increasing number of planes. But the index—itself admittedly inaccurate—moved erratically. The trend was a guess and extrapolation a gamble.

The various methods[1] used above were all employed by the Ministry of Aircraft Production and used to some extent to check each other. The other Supply Departments were, of course, faced with the same problem. Indeed the Ministry of Supply, Sir Oliver Franks's own department, encountered even greater difficulties because its output was still less homogeneous.

A final defect in this method of calculation may be noted. Total requirements were estimated and the figure so obtained was subtracted from the figure for the existing labour force in order to obtain requirements. Now these figures for present and required labour force were large relatively to the figure for requirements and a small error in either of them would result in a large error in the estimate of man-power required. There was no way of evading this difficulty.

Enough has been said to support the statement that the departments themselves were uncertain about the magnitude of their requirements. This was true of the Service Departments as well as of the Supply Departments. The Chiefs-of-Staff set down the number and types of warships, divisions, brigades, and squadrons they required, but there was no infallible formula by which these requirements could be turned into man-power. It was easy enough to say, for example, how many men were needed to man so many ships, but what allowance should be made for the large number of naval personnel permanently ashore? The Army had a long tail. Could it be docked? The numbers required for aircrews could be worked out, but the bulk of the Air Force was on the ground and the proportion of ground staff to a squadron was not invariable.[1] It is true that the services had approved establishment figures which often permitted a ready calculation to be made, but these were merely average ratios and were not necessarily above criticism. As for the heterogeneous requirements of the civilian industries, they were still more difficult to estimate and the arguments they provoked were often as vague as they were emphatic—an ominous portent for planning in peace-time.

---

[1] It may be asked why the man-hour coefficients were simply used as weights. Would it not have been possible to add these figures for all the aircraft in the programme and thus obtain a direct estimate of labour required? The difficulties may be enumerated briefly. (1) The man-hour coefficients related only to direct productive labour. It would therefore have been necessary to gross up these figures, und this was what in fact was done by using the figures as weights and relating total production in terms of man-hours to total man-power. (2) Figures for total man-hours were inadequate without some knowledge of the way in which these man-hours would be distributed through time. (3) Finally, anyone who feels inclined to be optimistic about such calculations should reflect on the complications introduced by differences in efficiency through time and as between one firm and another. In short, this method offers no escape from the difficulties described above.

In such circumstances it was only natural that departments should play for safety by putting forward a high figure which would safeguard their position if events turned out to be unfavourable. After all, each department had been charged with grave responsibilities on which the future of the country depended, and it was natural, and indeed right, that they should try to insure against risks. It was not for them to estimate the relative importance of their own needs as compared with those of other claimants. To begin with they had not sufficient knowledge, and the task was rightly left to the central planning staff. Apart from such risks, the departments tended to write up their estimates for another reason. Everyone was aware that there was not enough labour to go round and that cuts were inevitable. The cuts would not, of course, fall with the same severity on all departments, but few, if any, could hope to escape entirely. There was therefore a strong temptation to submit a somewhat larger figure which would contain a margin not only against risks in the field or the factory but also against subsequent cuts. We have noted this tendency already with regard to firms and it was indeed apparent throughout the whole war economy. The humble quarter-master sergeant was always on the scrounge for extra boots; far above him the Minister in his office was playing a similar game.

I do not wish to give an exaggerated impression of this difficulty or to suggest that departments set out deliberately to mislead the Cabinet. Their behaviour can only be understood if one bears constantly in mind that their own estimates were uncertain. A department might feel that an additional 175,000 was the best estimate it could make of its needs, but there was always the possibility that 200,000 would be genuinely required. In view of the danger of cuts it would then appear unwise to submit the

[1] A good example of these difficulties is afforded by the disagreements about the speed of demobilization. Cf. the debate in the House of Commons on Oct. 22, 1945, in particular Mr. Churchill's comments on the R.A.F.

lower figure, for the allocation might then be no more than 150,000, and if things turned out badly this would be 50,000 below what was needed. In such circumstances it was natural that the higher figure should be submitted and subsequently defended.

III

The estimates of requirements were first examined by a small committee of officials which included representatives of the Ministry of Labour and the Ministry of Production. The figures were carefully scrutinized and officials representing the claimants were invited to attend and justify their demands. In this way a certain amount of water was squeezed out of the estimates before they were submitted to the appropriate Ministerial Committee. The process of cross-questioning was then carried on at the higher level and a strenuous attempt was made to scale down the demands as fairly as possible by the extent required to bring them into equality with the anticipated supply of labour—itself a somewhat uncertain quantity. By the time the figures went to the Cabinet much progress had usually been made, but as a rule there was still some gap to be closed and a number of difficult issues had usually been raised which affected the nature and size of the war effort and the welfare of the civilian population to so serious an extent that the Cabinet alone could assume the responsibility of taking the final decision.

At this point the Prime Minister was frequently called on to play an important, and indeed indispensable, role. No one was prepared to yield, but there was not enough to go round and sharp cuts had to be made somewhere. Now it was not for the claimants themselves to suggest where these cuts should fall and provide a solution. The suggestion could only be made by someone who stood above the departmental strife and had the authority to issue a draft directive and insist that this should form the basis of discussion. It required the Prime Minister, and a very unusual Prime Minister at that—a leader with great strength of will and a capacity for inspiring affection and respect amongst his divided team, an expert on logistics with a remarkable ability for hard, quantitative argument and a great capacity for mastering complex reasoning relating to a host of difficult issues. It is to be hoped that Mr. Churchill's minutes will one day be published in full. An examination of those remarkable documents, which contain so much detailed information on an enormous range of topics, should destroy at last the strange belief that Mr. Churchill's wartime role was largely that of an inspired rhetorician—if indeed any presentation of evidence can remove so unaccountable a delusion.

The final struggle might thus sometimes take place on a draft directive from the Prime Minister suggesting a sufficient number of cuts to close the gap. The title of the document made it clear that a solution would be imposed on the departments whether they liked it or not, but the specific proposals it contained were only tentative and were often modified appreciably by the War Cabinet. It was always convenient to have definite cuts before the meeting at the outset, even if their author himself was uncertain about their wisdom, because in this way the departments could be forced to state their case again and provide fresh information. In short the discussions must have represented something between a scientific meeting and a public auction.

Thus in the end a decision was taken and the allocations made. The cuts in demands were frequently accepted without any change in the final programmes for the number of ships to be manned, aircraft to be produced, and so on. If a figure for requirements had been deliberately exaggerated in anticipation of future cuts, the cut was to that extent of no importance to the department; the minister might prophesy ruin to save his face but return quite cheerfully to his office. Otherwise departments were obliged to operate with a smaller margin against risks or were forced to make more strenuous efforts to economize man-power. In some cases, however, it was agreed that programmes would have to be reduced—for example, the bomber programme was scaled down on several occasions as strategy changed and other claimants received increased priority—and the planners had then to get down to work once more and examine the implications as precisely as possible. Meanwhile, however, the Cabinet allocations would be made on the basis of what information was available. It was rarely possible to postpone a decision until all the consequences had been fully examined.

Finally, it should be noted that these allocations dealt in the main with aggregate man-power and referred to the country as a whole. Of course it was often necessary to take account of particularly important complications, such as the supply of mobile women, and these were sometimes decisive in determining Cabinet decisions. But it was out of the question for a group of busy ministers to study the regional problems in detail or to spend much time in discussing the supply of different kinds of skilled labour. By and large it was left to the Ministry of Labour, together with interested departments, to determine the distribution of the allocations, and although it had at its disposal man-power controls which were far more stringent than any which can be used, or should be used, in peace-time, the performance of this task called for administrative efficiency of the highest order. That there were many failures no one would deny. If the figure for total unemployment was low, there was no doubt a certain amount of concealed unemployment, in the sense that some factories had more workers than they required.

So much for the man-power budget. It did not result in that nicely scientific distribution of the nation's labour force of which the 'planners' dream and, worse still, talk. The departments were far from certain about their requirements even when their programmes, themselves subject to serious error, were regarded as given. The neutral ministers and officials who were obliged to arbitrate between the conflicting claimants could not be expected to know as much about the needs of a department as the department in question knew itself. They might query certain figures and ask for further information, but where the issues were so complex and the scope for disagreement so great, they were liable to be beaten in argument. Such a defeat, of course, was rarely final. It did not mean that the victors escaped a cut, but it did mean that the cut was often somewhat arbitrary. The reduced allocation was frequently in conflict with the best information before the meeting, but since that information came from an interested party it was liable to be, if never inaccurate, at least one-sided. Everyone knew that departmental demands could be scaled down to some extent without seriously endangering programmes, and strenuous efforts were made to divine the water in the estimates. But in the end there was bound to be some margin of error. When the war was over a senior official, who had great experience in these matters, observed to me that he had long since seen the absurdity of the belief that allocations could be determined by means of scientific discussion with the departments. In the end, he said, the cuts must be more or less arbitrary.

In concluding this section it must be emphasized as strongly as possible that the defects in man-power planning were not the result of remediable faults in the war-time administration, nor can they be attributed to the lack of a sufficiently scientific approach. On the contrary, a most vigorous and sustained attempt was made to run the war as scientifically as possible. The Prime Minister himself appreciated to the full the importance of obtaining an adequate supply of statistics. He would not tolerate arguments which were vaguely expressed and unsupported by quantitative estimates, and his demands did much to improve departmental practice in this respect. He was supported in all this by one of the most remarkable teams of ministers in our history and by a civil service which contained within it far more ability than the country can normally spare from other occupations. The civil service has, of course, come in for much abuse and it is paradoxical that socialists are often i s most vehement detractors.[1] The permanent officials in particular are singled out for a great deal of dingy criticism. In fact, the permanents worked on cordial terms with a large body of temporaries drawn from all walks of life and sought the advice of the new-comers with great humility on the many technical points where they felt that their own training and experience had been inadequate. During these years the administration of the country was transformed. Great new ministries were created and an elaborate machine was set up for the central direction of economic affairs. The government statistical service was vastly improved and every effort was made to evaluate as carefully as possible competing claims on the nation's resources. These developments were not sabotaged by 'reactionary permanent officials'; on the contrary it was they in many cases who initiated the changes. The Treasury itself, that stronghold of 'unenlightened' officials. relinquished much of its traditional power and was content to play, in the main, a minor role throughout the war.

---

[1] Cf. the recent onslaught by Professor Ian Bowen in *Britain's Industrial Survival.*

It would be unnecessary to say all this but for the persistent tendency on the part of physical planners to attribute most of the defects of central planning to the folly and wickedness of those who did the job. This is in line with Russian practice. Since it is wicked to criticize the system, all faults must be attributed to sabotage and the offenders liquidated. It is some consolation that the sneering has not yet given place to shooting.

It would certainly have been possible to collect more information during the war and make more scientific allocations of man-power and raw materials. With regard to aircraft, for example, it would have been possible to learn a great deal more by employing several hundred engineers of the highest ability on this task alone. Unfortunately there would then have been very few engineers to make the aircraft, for in the industry men of such ability were numbered by the hundreds, not the thousands. The same is true in other fields. The facts of scarcity have at last been recognized with regard to industrial man-power and productive capacity, but the physical planners still make the assumption of a virtually unlimited supply of the factor of production which is so extremely scarce and in such inelastic supply—high-grade administrative ability.

The efficiency of the administration could, no doubt, have been improved during the war in spite of these limitations. It has been suggested that it could be improved at the present time by transferring some of the best officials from other departments to the Treasury which may be too small to cope with the enormous complexities of international finance. But the very fact that scarce administrative ability is wrongly distributed is merely a further indication of the unavoidable inefficiency of large units.

There is one other difficulty—familiar enough, yet often ignored—which must be mentioned in passing. Apart from the use of the pricing system, it is possible to delegate major decisions to only a limited extent. During the war an enormous range of topics came before the War Cabinet and a small group of senior officials. This was unavoidable because of the importance of the issues and because it was necessary to be familiar with other needs and problems in order to decide one of them. It is easy for an academic critic, who has spent twenty leisurely years in studying some specialized subject, to abuse these men for their comparative lack of knowledge in dealing with it and to forget the great versatility which they are required to display. What was true during the war is true to-day. The same versatility is required, and while it should excite our admiration, it implies superficiality and arbitrary decisions when it has to be carried too far. Furthermore, since 1940, some of the ministers and most of the officials have been working hours which, if suggested for industry, would lead at once to a general strike. Our administration is creaking under the strain and but for the great physical and mental endurance of a small group of men would have broken down before now. Sir Oliver Franks himself emphasizes the importance of this problem and the need for a solution. 'I am not', he adds, 'prepared to make suggestions on how it should be done.'[1] But it is surely a little unfair to raise a major issue in this way and then leave it unsolved without modifying his conclusions about the advisability of the central direction of economic affairs. He has in fact pointed to the final and insuperable obstacle to scientific, authoritarian planning. There are only two remedies—to put up with more or less serious inefficiency or to decentralize by means of the pricing mechanism.

IV

'Peace', says Sir Oliver Franks, 'is very different from war. . . . But the essence of planning and control by the State in relation to productive industry and commercial activity must be the same in war and peace.' That this is so cannot be doubted if what one has in mind is physical planning, but the moral to be drawn from war-time planning does not appear to be so encouraging as his lectures suggest. In addition to the difficulties mentioned above, it is difficult to see how adequate provision can be made for that expansion of the more efficient firms upon which economic progress depends as much as upon the hotly debated modernization of plant. It would be unwise to suggest that the pricing system offers an easy alternative, and both sides in the controversy have, perhaps, been inclined to criticize their rivals' thesis by ideal standards. Nevertheless, I am convinced that when the evidence is examined impartially it warrants the conclusion that a restored pricing system, coupled with drastic financial planning and associated with a few direct controls, offers more than the cumbersome mechanism of physical planning. No attempt, however, can be made to develop further the arguments in support of this conclusion in the present article.

UNIVERSITY COLLEGE,
OXFORD.

---

[1] Op. cit., p. 57.

A RECONSIDERATION OF THE THEORY OF THE INDIVIDUAL BUSINESS

By P. W. S. ANDREWS

1. Introduction

THIS paper puts forward the outlines of a theory of the individual business.[1] Its generalizations have been reached in the course of detailed investigations of individual businesses during an inquiry into the factors affecting the relative efficiency of small- and large-scale businesses. This research, which is still proceeding, was suggested initially by the late Mr. Samuel Courtauld while a Visiting Fellow of Nuffield College, and encouraged practically by his generous gift to the College. It would be inappropriate to let a footnote carry the burden of the author's debt to Mr. Courtauld or to the many business men who, by their generosity with their records, their time, and their patience, have made the research possible.

The theory with which this paper is concerned was originally arrived at piecemeal, in successive attempts to describe one business so as to compare it with another. Other writers have already put forward some of the propositions, even if only as possibilities, but the whole theory has been worked out independently in order to account for the business behaviour that I studied. The upshot of the theory certainly differs considerably from that of the theories of such economists as Kalecki. They think in terms of monopoly where I think in terms of competition, and I do not see the gross profit margin as a simple index of monopoly power.[2] To me, personally, the chief extraneous interest of the theory is that it tends to reinstate the Marshallian view of competition and to justify Marshall's ideas about normal price.

The theory of the individual business, which is so much the core of traditional analysis in this country, has become more important, rather than less, despite the recent preoccupation with the problems of social behaviour, the theory of aggregates, and the technique of social accounting. Its importance is shown by the contemporary discussion of industrial efficiency, of the impact of nationalization, and of the uncertainties affecting the relation between business behaviour and social strategy or between private and public investment. The theory of business needs restating in terms such that it can be conveniently incorporated into social accounting theory.

According to the views put forward here, the theory of the individual business has been frustrated by two things in particular. The first is the neglect of the existence of short-run reserve capacity (for no obvious reason in view of the writings of Sraffa, Clark, and Harrod), which appears to have led to errors in the theory of costs for an individual business in normal situations. The second is the neglect (to which Hall and Hitch have called attention)[1] to take time into account in the theory of demand. It is, in retrospect, surprising that economic theory should not have considered the long-run behaviour of demand, which is so clearly a factor in the business man's behaviour, whilst giving great attention to the theory of long-run costs, which, as I shall argue, is less important.

---

[1] A more detailed account of the theory will be given in a later book.

[2] Kalecki's analysis also seems to assume (constant) short-run elasticities of demand, which I deny as relevant, and do not need to suppose the business man to know.

This paper, therefore, falls naturally into two parts, the first dealing with the theory of costs, while the second discusses the important question of the determination of prices. The analysis is necessarily abstract, since the theory is intended to have general application, and it will not be possible to go very far into the detailed modifications that are necessary when one is looking at a particular individual business producing given products at a given time. The separate parts of the paper will each carry its own set of assumptions, but there is a general reservation which should be made before we turn to the detailed arguments.

When an individual business is being studied in isolation an adequate account of its behaviour can usually be given from two points of view—the market situation in which it is working and the organization which it has available in order to meet that situation. Any peculiar features of its history can usually be treated under these heads as they arise. But when one wishes to compare one business with another, it is usually necessary to introduce a third topic which can be left implicit in the earlier type of analysis, i.e. the relevant characteristics of the main personalities who have played their part in each business.

This is always a key element in the two factors of situation and organization which the observer has otherwise to take as given. Even when one would consider them to be of equal efficiency on balance, business men usually differ in the interest that they take in particular aspects of their functions, and their personal predilections and abilities leave their mark on the business. This personal equation is especially important in determining the size to which a given business has been allowed to grow or towards which it is straining. The attitude which the business man takes up affects both the market situation and the kind of organization which he evolves. To refer to another aspect of the same problem: while it is true that the immediate aim of the business man is to make profits, and that his business is where he gets his living, it is also true that it is his way of life, and that the profit-getting motive will be affected and moderated by his personal preferences as will any other man's income-getting motives.

---

[1] *Oxford Economic Papers*, Number 2, May 1939, 'Price Theory and Business Behaviour'.

This paper conforms to the usual theoretical custom of abstracting from the personal idiosyncrasies of business men, which is justified in the broad treatment of social phenomena in economic theory, but it is an important omission in the theory of the individual business as such.

2. Costs in the Individual Business

i. Assumptions

This section discusses the way costs of production in an individual business may be expected to change as the business changes its output. Cost curves, as drawn in current text-books, are thought to be misleading when applied in the study of individual manufacturing businesses. In particular, it will be asserted (1) that it will be normal for a manufacturing business to experience decreasing short-run costs; and that the rising branch of the short-period average-cost curve is irrelevant; (2) that the rising long-run cost curve, as usually drawn, is based upon a wrong conception of the way in which management factors affect costs.

The firm considered is imagined to be engaged in manufacturing as ordinarily understood, i.e. businesses engaged in extractive industries or in commerce are formally excluded, although the bulk of the conclusions will apply to undertakings engaged in commercial activities.

Secondly, the business is supposed to be manufacturing a single product of a homogeneous specification at any one time and will make only such changes in specification as are called for by the new production methods which naturally accompany any changes in the scale of production.[1] To make the business so rigidly single-product is indeed to take away the flavour of reality, but it would add unnecessary complexity to discuss a multi-product business, and it would not alter the broad conclusions.

Thirdly, the business will generally be assumed to be operating a single plant.

Fourthly, as a previous paragraph has indicated, we shall be concerned with *production* only, and shall not discuss selling costs.[2] Some assumptions will, however, be made about the markets for the firm's factors of production, since the prices of those factors will affect the costs of production.

ii. Short-run Costs of Production

Costs of production are of two classes—direct and overhead costs. The actual content of the terms cannot be clearly defined unless they are referred to the appropriate scale of output, but the distinction does more or less correspond to the prime costs and overhead costs of normal theory.[1]

---

[1] The qualification is necessary because it is usually true that one cannot make precisely the same product by large-scale methods of production as by small-scale methods. A product usually has to be redesigned for mass-production.

[2] A theory of the behaviour and consequences of selling costs will be given in my book.

All forms of direct costs have this in common, that they are directly incurred in, and can be directly related to, the production of each particular unit of product. They are of three kinds: (*a*) direct materials costs; (*b*) direct labour costs, which may include other expenses besides wages; (*c*) direct expenses, which are the other expenses of production which vary directly with the output of the product: an example is the running costs of power used for operating process machinery.[2]

Overhead costs comprise all other expenses of the business's operations. They include (*a*) overhead materials costs—the costs of materials not used directly in the production of the product and the use of which does not vary with changes in output; (*b*) non-process labour costs—the wages of all workers not actually employed on process work, such as yardmen, labourers employed for internal transport, &c.; (*c*) the wages of foremen and other supervisory grades of management;[3] (*d*) the wages and salaries of office staffs; (*e*) the salaries paid to management personnel; (*f*) rent, rates, and insurances; (*g*) the costs of repairs to machinery, renewals of loose tools, &c. The remaining type of overhead costs (*h*) consists of the 'financial' costs of writing down capital equipment on account of depreciation and obsolescence, so that the expenses laid out on the equipment shall all have been charged as costs by the time that it is expected to have ceased to be useful to the business. There is a big difference between this class of overheads and those listed earlier, since they carry no necessary obligation to pay out currently actual sums of money away from the business, while the others are all '*paying-out*' costs, and, unless covered by current earnings, will deplete cash balances.

When analysing the effects of changing outputs and sales on a business's financial position, it is useful to add to the strict costs the current obligations by way of interest payable for loans to the business. Failure to earn sufficient to cover them will not involve the same consequences as failure to cover the classes of overhead costs (*a*) to (*g*), but they will represent a current need for cash, and default may well cause a change in the ownership of the business or of some of its assets. In fact, *payment* of them may well be made in preference to providing for the 'financial' costs listed under (*h*).

---

[1] Direct costs are, in any case, the important part of prime costs from the point of view of price-determination, as will be seen in Part 3.

[2] It should be clear that, granted the assumption of a single-product business, it is not necessary to distinguish variable overheads; they have to be given special recognition only where it is necessary to distinguish between the output of a single product and the output from the plant as a whole.

[3] In small businesses, of course, some of these may also be engaged on process work, in which case their wages may more conveniently be reckoned in with direct costs.

Since overhead costs are not directly related to particular units of product, the way they are charged as costs will depend upon accounting procedure. The items (*a*) to (*g*) (and also the current interest obligations) will appear in the accounts as given amounts per period and, for our single-product business, may be assumed to be simply divided by the number of units of output produced in a given period, to give the appropriate average costs per unit. The amounts written off the value of capital equipment will be determined by an accounting procedure the exact basis of which may vary between businesses, and the speed at which the equipment is written down will, ideally, depend upon the probable life of the equipment—in practice, it depends upon the policy of the business man, the write-off having *some* basis in real cost but no necessary relations being statable. We may assume the business which we are discussing to employ a consistent policy, which will yield it a determinate sum per unit of time as chargeable to the accounts.

As already noted, it is necessary to make assumptions about the prices at which the business buys its factors of production. It will generally be assumed that they will not vary at all with changes in the size of the business's demand for them. Although no change in basic wage-rates will be allowed, it will be reasonable to assume that higher wages would be paid if the business produced a higher output by getting an existing labour-force to work overtime. This will mean that, for the period during which overtime is worked, the existing wage-rates will be increased by a definite proportion, such as one-quarter.

It will not be possible to discuss very fully the generally accepted theory of costs, which rests upon a Marshallian basis (modern usage, however, largely derives from the articles by Harrod and others during the famous *Economic Journal* controversy on costs, and from an article by Viner).[1] On the usual assumptions, the static law of diminishing returns is held to justify the short-run average-cost curve being drawn as U-shaped. The rising branch of the U, in particular, is justified by the assumption that there will be an optimum dosage of the direct cost factors, after which average direct costs rise and, in the end, more than counterbalance the effect of the falling overhead-cost curve.

However, the rising part of the average direct-cost curve, and hence of the average total-cost curve, even when it would exist, is not relevant to normal analysis. The normal situation is that the business man will plan to have reserve capacity, his average-cost curve falling for any outputs that he is likely to meet in practice, and his average direct costs, which the second part of this paper will treat as of crucial importance in the theory of pricing, normally being practically constant for very wide ranges of output.

---

[1] We would cite particularly the article by R. F. Harrod, 'Decreasing Returns', *Economic Journal*, Dec. 1931, and J. Viner, 'Cost Curves and Supply Curves', *Zeitschrift für National-Oekonomie*, 1931–2.

The present theory differs from accepted doctrine precisely in that it considers the constant level of average direct costs, and the reserve capacity on which it is based, to be *normal phenomena, irrespective of the degree of competition which the firm has to meet.* It has been urged previously that 'excess' capacity *may* be normal, but the claim has been put forward on the basis of the assumed effect of falling marginal revenue curves, and a consequent *falling* long-run cost curve. It is here argued that short-run costs will normally be falling even if the long-run cost curve is rising, and however competitive may be the market for the individual business.[1] In any case, the existence of short-run reserve (usually called 'excess') capacity has been ignored in most of the recent pricing theory. In Chamberlin's analysis, for example, the cost curves are clearly long-run.[2]

The reserve capacity which is so evident in the real world is not just long-run, as in Chamberlin, showing itself, in too many firms, each of which is too small for really efficient management. It is largely short-run in character, the manufacturing business having typically a capital equipment which could produce a larger output than it normally expects to produce so that its average costs would fall if it produced more from that equipment.

Such extra capacity has no necessary connexion with falling marginal revenue curves. A business man will always lay down his plant with a certain output of completed product in mind,[3] but, whatever the output, he will not invest in exactly the equipment which would theoretically produce just that output when fully utilized. As a general rule, the business man will not wish to produce at all near the theoretical capacity of his plant. Accidents happen to the best-looked-after equipment, and machinery will have to be repaired or renewed in any case. Many things, therefore, may cause a proportion of the plant to be idle at any one time, and a reserve, over and above the equipment which is strictly necessary for the planned output, will have to be kept to allow for these contingencies.

---

[1] This implies a rejection of the principle of tangency invoked by Professor Viner's draughtsman and others. That principle can be attacked on general grounds, and it is certainly not necessary for the long-run cost curve to be defined in such a way as to imply tangency with any short-run cost curve. But if the business man is assumed to plan for reserve capacity, the curves *cannot* be tangential, for the long-run cost curve will *necessarily* cut the falling short-run curve, whatever its slope.

[2] This neglect of the feature of short-run, 'excess', capacity is perhaps surprising in view of the original Marshallian discussion, the article by Sraffa (1926) which started off the *Economic Journal* controversy, and the plea by Harrod at that time that it may be normal.

[3] His choice of the method of producing that output will be considered when the behaviour of long-run costs is discussed.

Quite apart from such allowances for repairs and breakdowns, the output which a business man plans into his equipment will normally be greater than the output which he *expects* to produce when he lays down the plant. He will generally hope that in course of time, by accretion of goodwill[1] and continuity of service, he will face an expanding market. Further, he is producing in an uncertain world and will be liable to fluctuations of output. While he will be prepared to suffer his reverses, he will wish to take advantage of the occasional extra demands for his product. His capital equipment will be especially constricting unless he allows a margin in his plans—the margin being greater where the particular equipment is most likely to become a source of constraint.

The business man will, therefore, plan to have reserve capacity in order to get the elbow-room that he needs. He will also need some reserves in management personnel and overhead labour, but will not need so much, since emergencies can be met by working overtime, and since, subject to training time, his overhead personnel can be extended more easily.

In order to consider the effects of these reserves, it will be useful first to consider a relatively simple case, where the business man is operating a single process, e.g. where the plant is producing a chemical in one process, the sole nature of which is continuous chemical reactions between materials fed in in their raw state at one end and delivered chemically combined at the other. The only direct costs will be the costs of the raw materials—all other costs will be overhead, i.e. the cost of the building housing the process, the cost of the vessels or chambers in which the process is taking place, and the costs of the supervisory labour and management.

Total overhead costs, so far as plant and buildings are concerned, will be constant, and average overhead costs per unit of product will, therefore, fall hyperbolically as his output increases. His overhead labour costs will be constant in total for fluctuations in output not calling for overtime and will, therefore, fall on the average as output increases within that range. If overtime has to be worked, his labour overheads will rise by a fixed amount per production period and then remain constant. The average overhead labour cost curve will thus be on a higher level during the period of overtime, but will fall for increases of output.

Unless the plant is being operated at so low a level that some wastage of materials takes place (as it would in a chemical plant designed for a certain through-put of materials), or at so high a degree of capacity that the quality of the product cannot be so well controlled, the average direct costs may be expected to be invariable and constant, since the plant will use the specified quantities of raw materials per unit of product.

---

[1] This, of course, does apparently introduce 'monopolistic competition', but it is not to be taken as committing us to any belief in falling marginal revenue curves, as necessarily relevant to the analysis of such cases. That proposition is denied in the next part of this paper.

Diagram I shows the situation thus imagined for this simple case. Over the whole range of the capacity of the plant the business is shown as having constant average direct costs of production. The line showing the level of

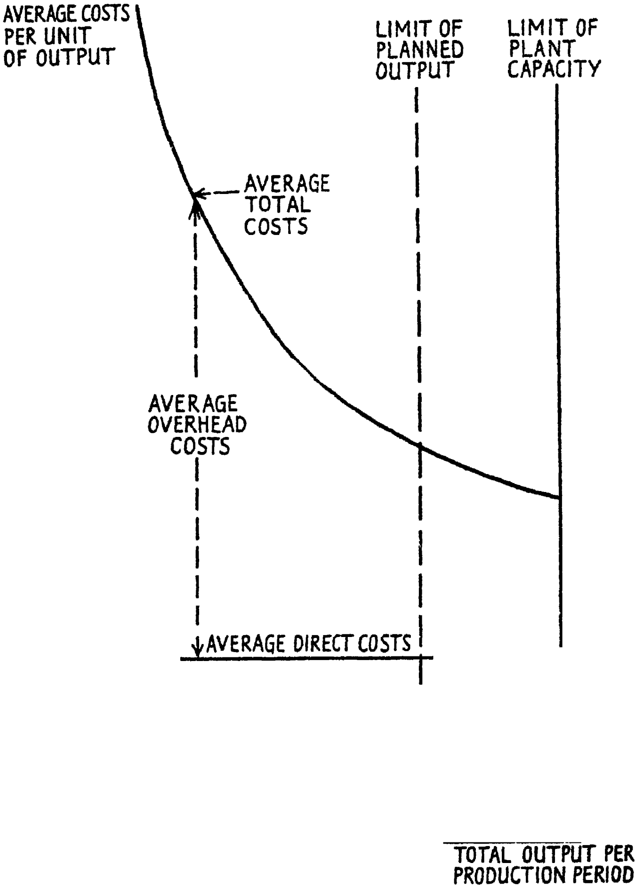

Diagram I

these costs is stopped abruptly because there will be a physical limit to the capacity of the plant. Average overhead costs are shown as falling continuously up to the same limit; the rise in the level of these costs due to working overtime has not been shown, since the difference would be small and it is desirable to keep this diagram simple. Diagram II shows the same situation in terms of total costs. This type of diagram is very useful for the analysis of the effects of varying outputs or prices on the profitability of a business. For that reason it tends to be used by business men in practice, e.g. in 'break-even' charts.

Most manufacturing businesses will, of course, be working in a more complicated situation. A typical manufacturing business will certainly have several process stages and some, usually an important part, of its labour-force will be engaged on process work, and will accordingly contribute to the direct costs.

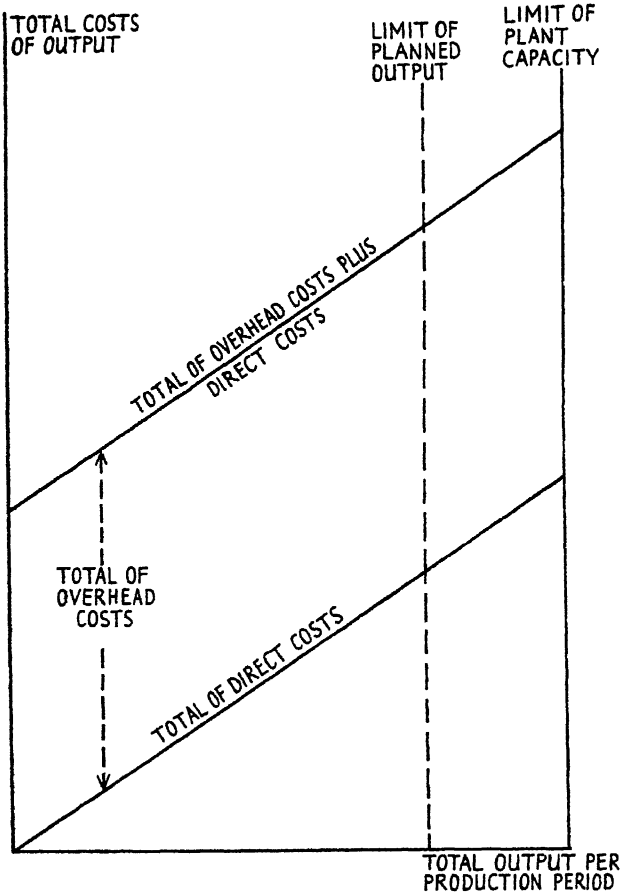

DIAGRAM II

When he is planning the equipment to produce a particular output of finished product, the business man will take each stage of production and assess the capacity of the machinery and equipment that is involved. From the present point of view the machinery may be ranged in order of the relation between the capacity of each unit-machine and the projected output. At the lowest level there will be the machines whose capacity will be the lowest proportion of the planned output; at the other will be the machinery or plant the individual items of which have the largest capacities in relation to that output. For each type of machine some reserve will have to be kept in hand to meet emergencies due to repairs. The repairs allowance may very well be largest for the small-capacity machines, since these are frequently the ones with the fastest moving parts and therefore most frequently need running repairs and maintenance attention.[1] The allowance to be made for possible increases in production will, however, be lowest here. First, a temporary but possibly large increase in output may be possible at the expense of deferring running repairs. Second, there will be a hidden reserve due to the possibility of working overtime. Third, deliveries of new units of such machinery can frequently be obtained without too long a delay. It will be here that the equipment of any industry will be most standardized, tending to be available as ready products from the machine-making industry. (For this reason, reserve capacity will usually be less where this equipment is rented.) Fourth, the installation of extra units of such machinery will not usually be a very great problem. Some allowance for the possibility can usually be made when the building is being designed. It will certainly not involve the major changes in layout, &c., that may be caused by additions to the bigger machinery. A business can therefore plan to develop in a more piecemeal fashion with this type of equipment.[2]

In the case of the most restrictive processes, the reserve capacity will have to be proportionately larger. To give one obvious instance, a business generating its own electrical power will naturally want to have a plant adequate for its needs—the economies of scale in the production of electricity are very large. But it will have to have a stand-by plant at least sufficient to cover a large proportion of its needs.

Further, it is frequently true of the more bulky heavier capital equipment that it has a relatively long life. The business man will, therefore, tend to look ahead and plan for reserve capacity here just because of the nuisance any limitation may become if his business grows. For example, he certainly will not wish to be installing fresh power-plant at frequent intervals with the waste that will occur from scrapping efficient plant which the normal development (and all business men hope for growth as a normal development!) has made too small.

The larger the output of a machine or plant the more likely it is that the business will plan to have reserve capacity. It will not be so easy to make piecemeal extensions. Such equipment tends very frequently to be made to the specification of the business, and additions to it frequently mean major reconstructions. These conclusions will still apply even if the machinery is worked on a continuous three-shift basis so as to get the most use of it before it becomes out of date or worn out. In that case, of course, overtime working is not available.[1]

---

[1] Of course, where this is so, the reserve may be partly in the form of spare parts.

[2] This is more likely to the extent that it is generally true that there is a tendency for whole departments to be dominated by machines of a particular class in terms of the relative importance of their capacities.

As output expands, apart from the relief obtained by working overtime, the business will approach the limits set by its equipment. The first limits—those set by the relatively small-capacity process machinery—may be met by extensions to equipment, each increment of which may be relatively minor. Up to this point, average overhead costs will be falling continuously. At this point, new machinery will cause the level of the overhead costs to rise. Ultimately, the business will run against the more fundamental limitation of its basic equipment and further output will not be possible without major reconstructions.

The way the other overhead factors are affected may now be considered. Labour is the most important of these from our point of view. At every stage of production there will be some supervisory labour, and the larger the relative capacities of the plant, the bigger the proportion of the labour-force that will, in effect, be overhead. If a large-capacity machine is working at only a part of its capacity, it will frequently require very much the same labour-force. Here, again, there will normally be some reserve capacity. Overtime may be worked, and people are also able to carry heavier responsibility for short periods of time than they can carry as a regular matter. For any short increase in output, then, it may be possible to carry on with the same supervision. If the output is sustained at the new level, then extra supervisors will have to be trained, but their absence will certainly not be the bottle-neck that a piece of machinery can be. For other grades of overhead labour the extra output may be carried on the basis of working overtime with the effect on costs analysed in the simple case.

The management of the business will normally be able to cope with increased responsibilities but may be augmented if those responsibilities become permanent.

Overhead costs thus are to be seen as fixed in total over broad stretches of output, and certainly over the range of output which the business expects as usual. It has not been necessary to make any significant change in the conclusions drawn from the simplified model. How about direct costs?

In the simple case, the only direct costs were raw materials. In the normal case, as has been said, the firm will have process labour. How will the cost of that vary with the output that is being produced? Can we assume that direct labour costs per unit of output will remain constant, subject to the usual qualification of a rise in the level of such costs for periods when overtime is being worked? We obviously can do so when labour is engaged on repetitive work and is being paid on a piece-rate basis. In many cases, however, labour is engaged on more varied duties, involving the application of individual skill, and so cannot be paid piece-rates. In other cases, for whatever reason, piece-rates are not paid, the labourer being on a time basis.

---

[1] In order to get the maximum degree of flexibility here the units of such large-scale machinery may be deliberately kept rather smaller than the technical optimum so that a larger number of units is employed. See page 74.

In the real world, of course, there are great differences in the skill and energy of individual workmen, but it is not a very unreal assumption that the actual performance differences between individual labourers on time-work in any one business will not be so great as the potentialities of the workmen, and for the present analysis it is suggested that it would be a legitimate simplification to assume that the *average* productivity of labour in the business is constant. It may be said without cynicism that, although new-comers to a business may display unwonted energy or lethargy, they will not long continue to work significantly away from the average when day-wages are being paid. If they do not come up to something near the normal working pace of the department in which they are engaged they will be dismissed; and if they work much harder, the normal forces operating inside a workshop will induce them to moderate their zeal.

So long, then, as a firm is varying output by working short time[1] or by taking on or standing off workpeople who are normally on its books, it may be expected that average efficiency of labour will remain unchanged. If a business were free always to discharge the least efficient workers, then a reduction in output might mean a fall in the average cost per unit of direct labour; but, in practice, businesses are usually compelled to share out short-period stoppages fairly evenly.[2]

As in the simplified model, therefore, average direct costs will be constant so long as it is not necessary to work overtime. The output will be limited, even when overtime is worked, by the capacity of the equipment, process equipment imposing a narrower limit than the less divisible machinery, &c. If at the earlier limit extra process machinery is acquired, this will cause average overhead costs to rise when the machinery is acquired, falling thereafter as they are spread over larger outputs. The effect of overtime will be that the (marginal) direct costs of the additional output will rise to the new constant level appropriate to overtime wages, the average direct-cost curve rising gradually to the marginal level. This increase in costs will, however, not affect pricing policy. Diagrams III and IV illustrate these conclusions.

---

[1] If a guaranteed week exists, then, of course, costs will be affected if short-time is worked.

[2] Of course, if the reduction in employment coincides with a general fall in employment in the area from which it draws its labour-force, then there may be a rise in the average intensity of effort of day labour—and the business may be more free to discharge the less efficient. There will then be a consequent fall in the level of the average direct-labour-cost curve. Equally, the existence of full employment in that area may cause a fall in labour productivity and a rise in the level of the average direct-labour-cost curve. This paper however, is restricted to considering changes in cost due to changes in the output of the individual firm.

iii. Long-run Costs

The theory of long-run costs is concerned with the way in which costs of production will change with changing outputs when the business is completely free to change its methods of production and to adapt them specially to the production of each particular output. The theory is therefore fully applicable only to the costs of a new firm constructing and operating a plant suitable for the output concerned. It will be convenient to restrict discussion to this case for the present. The case of a business which is already in being with a given organization and with a plant of a given size[1] is discussed in the appendix to this part of the paper.

It is usual to make a special assumption about the choice of methods of production, and the long-run cost curves of the text-books may be regarded as attempting to answer the question: What will be the cost of production of each output *when the business is so organized as to produce that output in the most efficient manner*?, and long-run costs so defined are, therefore, assumed to be the lowest costs at which the particular outputs can be produced. That assumption will be questioned later on,[2] but, given it, the theoretical cost curve is a reasonable concept. The business man could list various methods of production by which each output could be produced, and would be able to make forecasts of the costs of establishing and operating a plant using any of these methods.

The older texts refrain from making a single generalization about the shape of the long-run cost curve for an individual business. Three distinct possibilities are described—increasing, constant, or decreasing long-run costs. In more recent literature, the shape of the long-run cost curve is, however, brought within a single generalization, and such cost curves are usually drawn U-shaped with minimum average long-run costs at the point of optimum scale for the individual business.

---

[1] The establishment by such a business of a completely new and separate plant, although outside the scope of this paper so far, might seem sufficiently a new venture to fall within the theory that is being discussed, but decisions about the size and structure of such a plant would not be taken without reference to the plant already being operated and the conditions which are desirable in a complementary plant may well be quite different from those which would be preferred if the new plant were an independent entity.

[2] It has already been implicitly denied, see p. 59, n. 1.

The falling branch of the U is justified easily enough. There must be some outputs which, however efficiently they are produced, involve a plant which is on too small a scale to make use of available technical economies. Similarly, some sizes of a business will be too small for efficient management, the managers being too little specialized (e.g. engaged on process work and other non-managerial duties) or under-employed. It is the rising branch of the U that is more difficult to justify and, since that

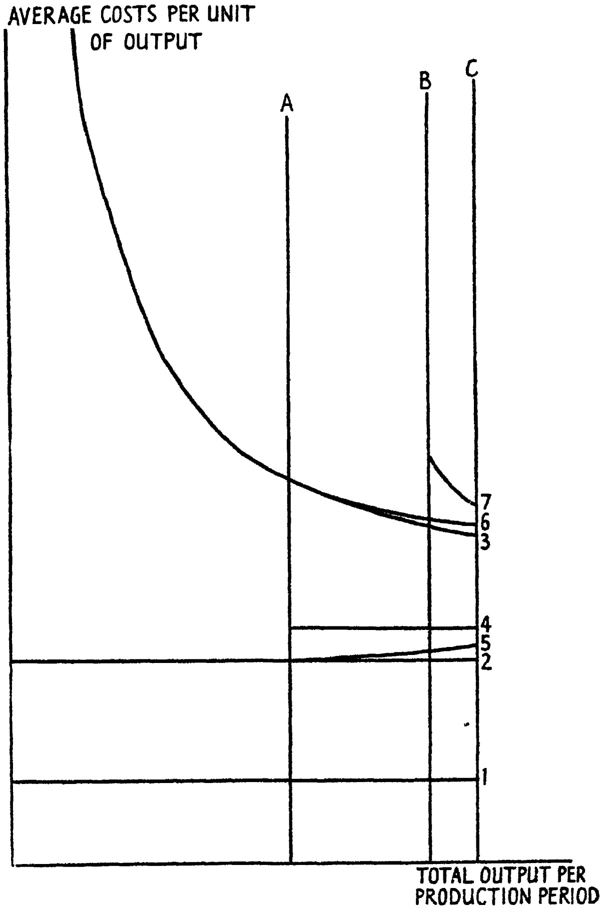

Diagram III

1 = Average direct material costs. 2 = Average direct costs with no overtime. 3 = Average total costs with no overtime. 4 = Marginal direct costs with overtime. 5 = Average direct costs with overtime. 6 = Average total costs with overtime. 7 = Average total costs with new process machinery working overtime.

A = Assumed limit to output from normal direct labour force. B = Assumed limit to output from process machinery. C = Assumed limit to output from machinery with greatest reserve factor.

is usually justified by reference to management conditions, let us look at that side of the question first.

Management's task is to run the business, supervising the non-managerial factors of production. A lowering or a raising of managerial efficiency may, accordingly, show itself *either* in a rise or fall in the level of these non-managerial factors per unit of product, *or* in a less or more than proportionate increase in the number and expense of managers. It seems

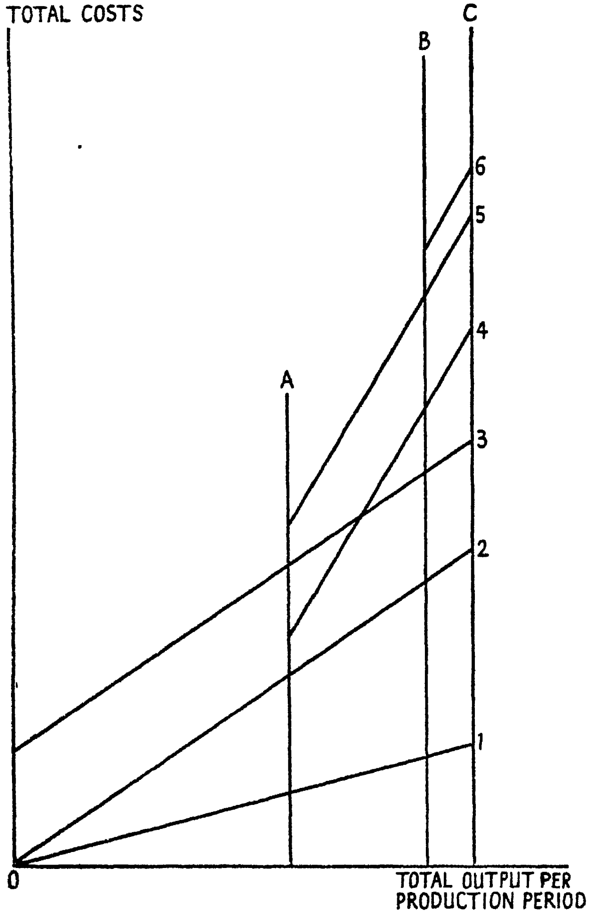

Diagram IV

1 = Total direct materials costs. 2 = Total direct costs, no overtime. 3 = Total costs with no overtime. 4 = Total direct costs with overtime. 5 = Total costs with overtime. 6 = Total costs after new process machinery and assuming overtime.

A = Assumed limit to output from normal direct labour force. B = Assumed limit to output from process machinery. C = Assumed limit to output from machinery with greatest reserve factor.

probable, in particular, that a lowering or raising of managerial efficiency will show itself particularly in a rise or fall in the level of direct costs of production, but that involves the short-run cost curves with which we are not now concerned. For the present purpose we can imagine a technical cost curve showing the average costs of production when the non-managerial factors of production are employed at the optimum level of managerial efficiency,[1] and can allow for changes in the efficiency of management by constructing a curve of average managerial costs, *it being understood that those costs may, in fact, show themselves elsewhere than in the costs of management proper.* The managerial cost curve is thus conceived as addable to the technical cost curve in order to get the effective cost curve.

As has been said already, the rising branch of the long-run cost curve is normally justified by management being regarded as, in the last resort, the one really indivisible factor of production, and, whilst the introduction of new personnel is admitted to allow of increasing specialization of *the* manager (who will become in the end simply *the* planner and chief sayer of 'yes' and 'no'), the fundamental indivisibility of his ultimate function is thought to remain. Accordingly, the law of non-proportional returns, to use Gide's term, is invoked to justify the concept of an optimum scale of business, so far as the management factor is concerned. Beyond this scale, diseconomies of management are thought to set in, and this will mean that, in the end, whatever the technical economies of scale, the cost curve must turn up and increasing costs be experienced for the business as a whole.

The generalization is frequently defended by implied appeals to 'experience', e.g. that, in fact, it is well known that large businesses are unwieldy and that there are decreasing returns to the factor of co-ordination, but the implicit assumption—that management will get continuously *and* progressively less efficient—should be questioned.[2]

Management should *not* be seen as a single factor combined in varying proportions with other factors and, therefore, subject to the simple law of non-proportional returns. There are *levels* of management. This is well known in everyday life, and it is possible to distinguish kinds of management techniques appropriate for different scales of output (varying with the type of business, i.e. in one industry the atmosphere of really large-scale management will be encountered at a far smaller size, measured either in value of output or in number of employed, than in another). Changes in the type of management will be called for only at discontinuous intervals, and it may reasonably be held that those intervals get farther apart with increasing scale. To give a crude example: in a particular industry there may be one system of management for a plant with under 100 workers; the next system of management technique may be applicable to the scale from, say, 90 to 500 workers, the next from 400 to 1,200 employees, and the growth in scale may then involve a system which is applicable to sizes from, say, 1,000 to 5,000, and so on. At each critical point in the scale the management will have the choice of being stretched to the limit or adapting itself to the new larger-scale technique (which will mean adding some personnel which will be overhead for the new range of scale—thus introducing one factor, however small its relative importance, which will make for decreasing costs over the new scale). Whatever the degree of inefficiency at any particular level of managerial scale, the business will probably be able to carry on without any notable decrease in average efficiency until near to the limit of its system of management, when it adopts the technique of management appropriate to the larger scale on which it is entering.

---

[1] The costs being calculated, e.g., as such estimates usually are, with the help of assumed figures for the capacities and performances of the various factors of production. A correspondent with wide experience of business projects says: 'I have seen a good many estimates of "technical long-run cost curves" (though not so described) offered as a reason for providing finance for large-scale reorganization or mergers. They usually turned out to be the less convincing kind of fairy-tale.' I quote this to emphasize that such a curve is a theoretical concept and to reinforce my later statement that, in fact, no business succeeds in producing at those costs—i.e. no business is 100 per cent. efficient.

[2] It has to my knowledge been questioned by Harrod and by Kalecki, and Kaldor has pointed out that the usually assumed fall in managerial efficiency with scale could hold only under dynamic conditions.

The way the efficiency of business varies as between the various levels of management must differ from industry to industry (and particularly with complexities arising on the selling side—outside the scope of this part of the paper). Management may well continue to get progressively less efficient at each higher level. Equally it *may* not, and we would put forward as a fairly plausible proposition that, *even* if we accept the usually assumed decrease in managerial efficiency, the rate of decrease will slow down with successive levels of management. *Still accepting the text-books' generalizations* then, we suggest that the 'management' average cost curve should be seen as a series of plateaux, each wider than its predecessor and each at a lesser height above its predecessor than that was in turn above its, the rise being steepest when the business gets too large for the personal management of the business by a single manager—this particular limit varying significantly according to the industry. If it *is* producing a *single* product, the business must have been so much organized already, at some scale, that larger scale will be simply a matter of routine. The management would then be able to carry on reasonably near to its existing level of efficiency. There is a general feeling that, at some very large scale, a business becomes so much of a routine that it suffers from not being able to adapt itself quickly to the vagaries of everyday life.[1]

---

[1] We have heard this alleged to be the case for one very large British manufacturing business, but we do not know the facts. On the other hand, it should be said that there ought to be some economies available to this very large business—e.g. through the employment of specialists, and that, for this reason, the foresight of the business may well be more

Our generalization, then, excludes, by reason of ignorance, the business which is so large that it becomes more akin to Civil Service than to business management, when the common opinion is that it would be relatively inefficient, even though the Civil Service has admittedly developed efficient techniques for managing organizations which are too large to be efficient!

As a matter of fact, firms appear to come up against the problem of systems of management much more usually when there is a change in the management due to the passing of the persons who brought them to the scale at which they are operating than through any simple further growth whilst those persons remain in charge. A person managing a business of a given size has usually a good idea of how to run it as efficiently at a *relatively* larger size. In any case, he sees the business as his child, and will have at least that reason for being an efficient manager and, in addition, will have built up his technique around himself. His successors always find it difficult to adopt the technique that was personal to their predecessor and frequently adopt the technique that would be appropriate to management on a still larger scale to simplify the problems of the changeover and to reduce the management more nearly to a routine.

So far as *technical* factors are concerned, it is usually assumed that there is an optimum scale of plant, this being justified by examples of businesses employing factors of production with *given* indivisibilities, e.g. employing units of machinery and plant of *given* capacities. The optimum size is then that of the plant producing an output which is the least common multiple of the various machine-capacities.[1] In terms of this concept, technical factors imply no more than constant long-run average costs of production once the optimum scale of plant has been reached, so long as the increase in scale takes place by duplicating the optimum combination of factors.

If the growth in scale is by a smaller amount, then, of course, average costs will rise until it is again possible to utilize fully an optimum unit. But this undulating curve, as Stigler calls it,[2] cannot have the rising trend required by the theoretical curve. Further, at a sufficiently large scale, the finite increase in scale necessary to get once more to the optimum combination of factors can be treated as relatively so small as not to impair the continuity of the curve, in which case the average technical cost curve

---

efficient than it could be for a business not of its large scale. One may cite the example of the employment of professional economists—surely profitable on a full-time basis only for a very large firm, but certainly profitable for a business whose output is a significant proportion either of the National Income or of the world output of its product, but it may, of course, be said that the employment of such experts will bring economies only to businesses engaged in a much more uncertain and complex world than that postulated for the cost curve of a single-product business under static conditions!

[1] E.g. H. A. Silverman, 'The Optimum Firm in the Boot and Shoe Industry', *Oxford Economic Papers*, Number 6, 1942.

[2] Stigler, *The Theory of Price*, 1947.

could be regarded as reasonably approximating to the conditions of constant costs.

The previous two paragraphs have accepted the assumption that there *is* an optimum size of plant from the technical point of view. It is difficult to accept the assumption, however, especially for the single-product plant, that growth means merely an enlarged plant, which is just a recombination of the machinery and equipment used at the smaller scale. A continuous increase of scale must always carry further technical economies at *some* point.

The larger the output going through the works, the more likely it will be that some process or other could be further mechanized. The introduction and elaboration of mechanical methods in ancillary operations such as internal transport is another possibility. Again, economies in maintenance and repair are likely to be always available to some extent with an increase in scale. Whatever the scale of the plant, maintenance and repair will necessarily be carried out on a smaller scale, and a large part will frequently be in the handicraft stage. Increasing scale, by increasing the runs of repair jobs of particular kinds, will enable some further economies to be achieved. There seems no reason to suggest that costs of maintenance do anything but fall continuously (although probably not sharply) with increasing scale.

Further, the single-product assumption relates to the end-product; the business is certainly not to be restricted to a single process, and continued growth may mean that it will become more attractive to manufacture for itself some of the goods that it previously purchased, when the technical economies of scale in that line of business will be added to it.[1]

Apart from the management factor, still assuming that the prices of the factors of production are not affected by the growth of the firm, it is, therefore, contended that the long-run cost curve should always be seen as falling with increasing scale. A rider should be added, however—that the larger the scale already, the less important *relatively* will be the incidence of further technical economies.[2] The fundamental redesign of equipment, or changes in layout and methods of production, will give way to the relatively less important development of mechanization of parts of the individual processes, &c. The more spectacular gains from increased scale will be exhausted first, and the technical economies of a still larger scale frequently consist of the further development of what has been done already. The reduction in the relative strength of technical economies as the plant grows must mean that, at a relatively large scale, the further economies available must seem of greatly diminished significance.[1]

---

[1] Of the same kind, perhaps, are the gains from undertaking its own research, both that which is simply devoted to the more effective control of the business's processes and the more fundamental kinds. In this field, economies of scale are notoriously likely always to be present, although in fundamental research they are perhaps not relevant to the single-product business of our assumption. Such gains will, strictly, affect the *level* of the cost curve from the point of scale at which they operate.

[2] The weakening of the force of the technical factors making for the expansion of the size of the plant is one reason why conditions in the labour market will tend to induce large businesses to operate separate plants in different labour markets. Where a business is producing many products, of course, the fact that economies of scale will first arise in, and be mainly available in, the common processes or services will be a restraint on any tendency to split up the plant, but, at a certain size, the possibility of economies through turning a plant entirely over to one product or to a more restricted range of product may be disintegrative.

But there is one further criticism of the theoretical long-run cost curve—the assumption that, for any given output, the business man will choose the method of organization that will enable him to produce at least costs must be questioned. The necessity for strategical reserves will affect the long-run cost curve as well as the short-run and will show itself in the nature of the business man's long-range plans and the general lines of his productive organization. This will make him seek for flexibility in order to be able to meet any likely need of varying his output (and his product). It will therefore not cause only short-run reserve capacity, but will also affect the type of organization and, therefore, the *types* of overhead factors of production employed. For example, in choosing managers he will seek for men with flexible minds rather than those relatively better at routine (i.e. whose efficiency is greater *given* the size and nature of the output to be produced). Similarly, in the choice of maintenance labour he will prefer a craftsman who is good at turning his hand to anything rather than one who will be better if his work is completely a matter of routine. But he will be *especially* restricted in his freedom to change his plans in future by his choice in the matter of investment in equipment. Although the short-run reserve capacity that is postulated will mean that the capital investment will be pushed farther than necessary for the particular output that is expected, it will not be pushed so *deeply* as if it were certain that neither product nor output would be varied.[2] It follows that, in terms of the short-run cost curve, average overhead costs *and* average direct costs will each be higher than they would be in a less uncertain world, and if there were all the infinity of choice postulated by the normal cost curves.

The effects of this will be particularly marked in the relatively more capitalistic industries (in the Austrian sense of employing more roundabout processes) or those which have the greater expectation of fluctuations in output. Reserve capacity is, therefore, most marked in the capital goods industries. *Within* an individual business, the effects will be most marked in that part of the plant where the units of equipment are largest in size relatively to turnover. Each unit of equipment will tend to be organized on a smaller scale than would be ideal for the output that is planned, in order to get the flexibility of a larger number of smaller units *and* to reduce the minimum desirable excess capacity. Where output is especially liable to fluctuate, particularly indivisible processes may be operated by, say, four or five units of equipment, when costs could be reduced if they were replaced by a single unit taking care of the whole through-put. Such an elephant would be economical only when running relatively near to capacity; it would have to be duplicated to some extent to get reserve capacity, and its operation would be seriously less efficient than the smaller units as the through-put declines—which situation could be better met by completely closing down one or more of the smaller units that are the alternative.

---

[1] Especially when taken alongside the effects of the short-run excess capacity which this paper stresses as being so much a normal occurrence.

[2] So far as the first of these factors is concerned, the author has several times had the instructive experience of visiting factories, one of which was run by a producer facing the normal risks of the particular product's market, and the other of which was run by a business that was really operating in the next stage. It was frequently the case that the latter business would deliberately plan for a factory considerably smaller than its usual needs, so that this could be operated full out with all available economies. The difference in the efficiency of the two types of factories, as measured by costs of production and as illustrated by the methods of production, has to be seen to be believed.

The desire for flexibility will also mean that ancillary processes such as internal transport will be far less automatic than they might be. Conveyors and such-like equipment will always be used relatively sparingly, since they are frequently sunk in a particular manner of producing a particular product.[1]

The removal of the least-cost assumption clearly makes it still more likely that technical costs of production will fall with increased scale of production—and, therefore, that they will be available, to some extent, to counteract any decreased efficiency in management, especially since they may well operate continuously, whereas management has been thought of as likely to become less efficient only at critical points in its scale, maintaining its efficiency over the intervening range.

Whereas Chamberlin's long-run 'excess' capacity appears to consist of management which is not fully extended at the given scale, we should expect that in the real world reserve capacity will exist in capital and other overhead equipment. This type of reserve capacity is 'excess' only in terms of theoretical criteria based on the postulate of the minimization of current costs of production, and it has been argued that it does not involve waste in the fluctuating conditions of the real world, nor does it necessarily result from monopoly.

The shape of the long-run cost curve will be the result of adding together the average managerial cost curve and the curve of average technical cost per unit of product. It is certainly not possible to make any very precise statement about its over-all shape. The technical cost curve should fall continuously,[1] although it will probably get less steep as scale increases. The managerial cost curve will fall for part of the range, and will probably rise again, at least for a time, after which it may or may not level off. What happens to management as scale increases is thus very much a matter of conjecture and will probably vary from industry to industry. Even *if* management should continue to get less efficient with increasing scale, that will only be as between levels of management, and the consequent rise in managerial costs will probably diminish as the scale of management rises. Two very broad conclusions, however, do follow: (1) that at a very large scale of output the business will get as near as matters to the condition of constant long-run costs of production—if we had not accepted the assumption that management becomes less efficient for all increases in management scale, then the weight of the technical factors would lead to decreasing long-run costs once management was at the routine level of efficiency; (2) that over any given scale of management the business will be in a condition of decreasing long-run costs.

---

[1] This effect will, of course, operate especially for the multi-product plant with which we are not now concerned.

It may be noted that, where diseconomies of management are most likely,[2] e.g. for a product where, for whatever reason, the processes are largely in the hands of hand-craftsmen (this will include industries where the machines are really power-driven tools), or, again, where the costs of the raw material are a large part of the total costs, so that continuous supervision of its use may be necessary, economies of scale are likely to be less important, because other circumstances limit the degree of mechanization. Conversely, in industries where technical innovations come most easily—where mass-production methods can be used—the sheer production side of management is most amenable to routine methods, and diseconomies to management are likely to be less important. Accordingly, a business will be more likely to be in a position of decreasing long-run costs the larger is the typical firm in its industry. But, however much costs may decrease with increasing scale, we should still expect that long-run economies of scale will be relatively unimportant compared with the falling slope of the short-run cost curve.

It will be seen that firm generalization cannot be made about the over-all shape of the curve of long-run marginal costs. It will fall over the early stages in the growth of a business; over any level of management, it will be below the long-run average cost curve, possibly rising at the limit of one level of management and falling at the commencement of the next; and it will approximate to the long-run average cost curve for very large scales, when the business will be approximately in a condition of constant long-run costs. Much more important in our view, however, is the curve of marginal long-run technical costs appropriate to the business which is under consideration. This curve will never rise above average technical costs; it will generally be falling, although it may have stretches where it is more or less constant. In actual conditions it will probably not tend to be very low in relation to the curve of average technical costs.

---

[1] Its fall may become sharper at critical points where the business becomes large enough to embark on new kinds of technical economies.

[2] It must be remembered that this part of the paper does not discuss the selling side of the business at all, and ignores the fact that limitation of the market for one product may frequently lead to a business adding to the complexity of its management by becoming multi-product.

It is difficult to discuss in the abstract the way changes in scale, i.e. along the long-run cost curve, will affect the short-run situation in which the business man will find himself, but a few generalizations may be attempted. If businesses employed the same technical processes at all sizes, the slope of the short-run cost curve should diminish with growth of scale, since, thanks to the reduced importance of indivisibility, the relative degree of reserve capacity would diminish, but, as we have already argued, growth will probably mean an intensification in the use of capital and other overhead factors. Accordingly, the degree of reserve capacity may well increase, and, therefore, the slope of the short-run cost curve may be steeper for larger sizes of the business. Certainly, in practice, in industries where economies of scale are important, the reserve capacity of the bigger business is often greater than that at which smaller businesses operate, but marketing strategy plays its part in this situation.

Restricting ourselves to technical factors and to the effects of technical economies available only to plants of larger sizes, capital and other overhead factors of production are characteristically substituted for direct factors of production (particularly for direct labour), and a growth in scale accordingly will usually mean a fall in the level of the short-run average direct-cost curve. A large-scale business will, therefore, usually have lower average direct costs per unit than will a smaller-scale business, even if its average total costs are the same. This means that it will be more easily able to survive in any period of low prices.

APPENDIX: SOME OBSERVATIONS ON COSTS IN PRACTICE

This appendix will consider some aspects of the behaviour of costs of production as they are encountered in practical analyses; the heading refers to its contents as 'observations' to emphasize the fact that systematic generalization of the kind that has been made so far will not be possible, which is why the normal theoretical discussion stops where it does. In the actual examination of an existing business it need not be left there, since the data usually make it possible to decide what factors should receive special attention and what should be ignored, and it is frequently possible to assign limits to the strength of the factors in which one is interested. Certainly one can, in practice, usually maintain a level of generalization which is appropriate to the business. Here, in a general paper, one can make only general observations.

The assumption of a constant price-level for *all* factors of production is one which must, to some extent, be unrealistic. Although it may be tolerably acceptable for any change in output which is appropriate for a given short-run cost curve, it is more questionable for changes in scale, since *some* change could be imagined for any business which would make that business so large that it would have an appreciable share of the market for at least one of its factors of production, and, in that case, the price of that factor could not be assumed *a priori* to be unaffected by the size of the business.

With present knowledge, little general correction can be made. So far as concerns raw materials in wide industrial consumption relative to the size of the business, however, it seems that, perhaps contrary to received opinion, purchasing economies proper are likely to be unimportant over quite a wide range of scale as compared with technical economies. In one particular case, which seemed as likely to afford purchasing economies as could be imagined, they existed, but only for so great a disparity of business size as made it certain that so small a gain was not worth taking into practical account in the factors accounting for the growth of the large business, and they were certainly negligible against the economies of scale on the technical side.[1] To the extent to which such economies exist, of course, they *will* reinforce the effects of technical economies and make it more likely that the cost curve will fall.

The labour market is, perhaps, the one where firmer generalizations can be made, as it is the one where the effects of increasing scale are most important, especially since they will go against the weight of the generalizations which have been made on the basis of the usual fixed-price assumptions. In the short-run, it is nearer the truth to say that a business has a unique labour-market than to say that it shares its labour market with other businesses. The established business will have relations with a group of workmen whom it normally employs, and who normally look to it for employment, even though, at any one time, they may be temporarily unemployed. These will be available to it without the standard rate of wages being affected, and the degree of their utilization may be increased, as has already been said, by overtime—which will affect the levels of the cost curves but will not greatly affect their slopes. In the short-run, then, a business will increase its output by overtime, from its normal labour-force, and a growing business may work on an overtime basis for the whole period during which it is growing.

There will be a more floating body of persons who are not so clearly attached to any one firm, and who may be attracted if they are offered the prospect of full-time employment.[2] These again will come to it gradually without the standard rate of wages for the business necessarily being altered. So far as unskilled labour is concerned, in large industrial areas such as that around Birmingham, there may be a relatively large pool of labour of this kind normally available, if time is given for them to get to know the prospects. *If* the business wishes to expand beyond this or at a faster rate than the extra labourers come to it at its existing wages, it will have to bid away workmen who have a preference for remaining where they are already. Accordingly, the further we push our long-run cost curve *during a given interval of*

---

[1] It is frequently forgotten in discussing purchasing economies that some are available only to the small business. A small engineering business, for example, can well make use of occasional bargains in the second-hand machinery market, refurbishing them for its own service, but it would not pay a large business to do so, since the small lots that turn up would not be of much use; another example is the small business in the boot and shoe industry, which can similarly make use of odd parcels of leather.

[2] The existence of overtime, by the way, is never a deterrent to new-comers in ordinary times. In its own interest, the business will keep it within reasonable limits, and regular overtime is some sort of guarantee of future prospects, being regarded as one of the signs of a business which is vigorously growing.

time, the more likely it will be that there will be an upward pressure on the business's wage-rates. But it does not follow that a business will let itself get into this position. First it will have good reasons against mere wage-competition; and these will not be the simple monopsonistic reasons that may naïvely be adduced. As a continuing business, it has an interest in settled relations with its workpeople, and too great a rate of growth, especially if induced by the wage-packet, is liable to bring disturbing factors here. A continuing business will wish the new-comers to settle down in the business and cannot be so sure of this if it makes too rapid a development or if it gets its labour by snatching tactics.[1] Taking the cost curve as it is usually described, there must be some scale beyond which the growth of a business in a given time period must mean a rising wage-level, which, in turn, must cause the long-run cost curve to rise if still larger scales are achieved in the same period. We do not think that this part of the cost curve has any relevance to practical conditions. Growth does not usually come in such great spurts relative to the existing size of the business, and in the rare case that it does the business usually becomes multi-factory, locating an additional plant in some other area where it will have a new chance to grow.

If the normal idea of the long-run curve is retained, it must be seen as non-reversible. A declining business will be left with at least some equipment and organization which is too large for its short-run output. Accordingly, its long-run cost curve will lie above that which was appropriate to that output during its previous growth. This case is relevant only in the analysis of the pathology of business strategy, and would take the argument into wider questions.

When the actual behaviour and experience of a given business in a given market situation is being considered, it is, of course, the short-run cost curve that the business man sees as immediately relevant and which he acts on. In analysis, therefore, and particularly with reference to practical pricing policy, it should be our starting-point. Given the data, at least a working approximation of the short-run cost curves is easily constructed. At any given time a fairly satisfactory division can be made between overhead costs and direct costs (most businesses are multi-product, and variable overheads present a little difficulty, but allowance for them can usually be made with a fair degree of approximation). The direct costs can then be assumed to remain at their current level, and the total overheads can be taken as constant. From this, one can construct fairly reliable estimates of the effect on the business of normal fluctuations in output, i.e. those allowed for in the organization of the business. In rare cases more elaborate statistical methods will enable us to make better estimates. A good example of these is given in the evidence presented by the United States Steel Corporation to the Temporary National Economic Committee. It may be observed that the statistical curves conform to our generalizations, since a straight line gives a good fit to total costs over the period, as estimated free from entangling considerations such as the changes in factor-prices. It follows that marginal direct costs are estimated to have been constant for the very large fluctuation of output over the period covered by the statistical study. However, the kind of procedure that has been outlined frequently serves the business man in his policy-determination.

---

[1] It is not possible, without stating detailed cases, to consider the effects of the fact that a business does not work in isolation—that, quite apart from circumstances which are peculiar to it, it will share many events in common with other businesses who are working in its product-market or who share its factor-markets. The labour market does, however, lend itself to at least one simple generalization. The trade cycle will generally exercise a depressing or an expansionist influence on all businesses simultaneously. In any case, granted the fact of the localization of industry, a business will frequently be sharing its labour market with other businesses producing similar products, and those will certainly have similar experiences over the trade cycle. It follows that all the businesses in a particular labour market will be wishing to expand their output at the same time. When there is such a general pressure on the market for labour of a particular sort, the course of wage-rates will be upward, even for the existing labour-force. Once a certain degree of full employment has been reached, wages will start to rise, and increased pressure to get output will cause them to rise still farther. In the trade cycle, therefore, direct costs will rise with the rise in general output. It is, perhaps, for this reason that the normal short-run cost curves have appeared to be useful weapons for the social analysis for which they are used, but that is no justification for their use in the analysis of the behaviour of individual businesses.

We are now concerned with how the actual long period looks from the starting-point of the short period, and how it differs from the theoretical concepts. As already indicated, the hypothetical curve of long-run costs is of nothing like the importance that it is in the broader analysis of social conditions for which it is used in the text-books. However, the business man will usually have at least a hazy idea of the opportunities open to him from continued expansion, but they *will* be hazy, and the haziness will increase with the time that it would take to make a given increase in scale.

Looking at it from the business man's point of view, suppose that current output gets to the limit of the usual range of fluctuations, and that the pressure on output has come consistently enough to look like persisting. Some adjustments will have to be made, and, as already indicated, one of the first effects will be the working overtime of personnel—this will usually occur before the limit of the plant has been reached, overtime being worked in the 'bottle-neck' departments. That will mean a rise in the level of the direct-cost curve for the enlarged output which is being produced.[1]

If the high output persists, then the business man will start to make more permanent adjustments. He will be willing to take on and train extra labour, and, since this will mean working his process machinery more fully, he will probably order extensions here—which, assuming the growth only of the particular business with which we are concerned, will usually be possible without too much delay in normal times. The probable pressure on his labour supply will also make him regard with favour the deepening of his capital structure, using equipment to eke out labour. He will certainly be guided by his knowledge of, or guess at, his long-period technical cost curve, the general properties of which have already been discussed. It is worth stressing that *it is his technical cost curve which is relevant to the planning of the business man.* No business man would *expect* to be a less efficient manager at a larger scale; he will tacitly assume that he will remain as efficient as he is for any increase in scale that he is likely to make. Realized costs may, therefore, differ from the expected costs on which he plans, but that cannot affect his planning. It follows that to a business man long-run costs will usually be expected to fall with the growth of the business, and, therefore, that rising long-run average costs are ruled out of the analysis of business strategy, though they *may* be relevant to the analysis of a business's history.

Time is of the essence of the business man's thinking. There will be a limit to the rate at which he will be prepared to grow. He normally views a new scale as a position which once taken calls for consolidation. He will expect his costs at first to be higher than they will become when he has achieved that consolidation—in fact, for any given scale, he will expect, other things being equal, that his cost curves will fall over time, that he will always be able to make some improvements in the light of experience. The idea of an optimum size of business is outside his usual way of thinking. A reasonably efficient management will always expect to be able to manage a bigger business at lower costs if it is given due time in which to make adjustments, and it will take all the time that is necessary.[2]

---

[1] Such a rise in costs will not affect pricing-policy as a rule. Overtime will be seen as a temporary expedient, and granted the existence of the overhead costs, the larger output will still be profitable without disturbing the price-situation, even at the high direct costs (i.e. labour will normally be costed up at non-overtime rates).

[2] In our experience business men take a good deal of time to make long-run decisions such as are entailed in major extensions. They therefore try to make them in advance of their becoming urgent.

One final observation may be of interest. We have discussed the provision of an empirical content to the short-period cost curve; how far is it possible to do the same for the long-period curve? In some industries, where technical reasons make it possible to convert even multi-product output into equivalent units of a standard product, it is possible to use expert advice to construct a hypothetical technical long-run cost curve[1] which may be used as a criterion against which to assess the degree to which an existing business achieves a degree of efficiency which is thus thought to be *potentially* open to a business of its particular size, on the basis of the most up-to-date technical information.[2] Such a curve, of course, will extend over a far wider range than will be relevant to the planning of any particular business. The business man will usually have a good enough idea of the shape of his technical long-run cost curve in the neighbourhood of his existing size, but that will be only the beginning so far as he is concerned. Non-technical factors affecting efficiency frequently turn out to be of even greater importance than technical considerations, even for industries where technical economies seem spectacular relatively to other industries. These other factors can only be analysed in the light of the details of the business's marketing strategy, and of the development of its market, but, where the empirical technical cost curve can be constructed, it provides a help in assessing the relative strength of the other factors, and as a starting-point for the detailed questions which will uncover some of those factors, even if only in a qualitative way.

3. The Determination of Prices

i. Introduction and General Assumptions

This part of the paper gives the outline of the theory of the normal price-policy of an individual business; a general analysis of situations in which that policy breaks down is reserved for later publication.

The assumptions and reservations which will generally be made are as follows:

(1) As in Part 2, the business is assumed to be a manufacturer.
(2) The analysis will be restricted to the product-market.
(3) The business is first assumed to be single-product.
(4) The conclusions of the earlier part will be taken over as assumptions —i.e. (*a*) that average direct costs are constant for all outputs, (*b*) that since long-run pricing policy can be based only on the expectation of business men, average long-run costs of production will, at worst, remain constant, but will probably decrease at a slackening rate with the growth of a business. For this reason the business man will use his existing (short-run) costs as a basis for his long-term pricing policy.

---

[1] But see p. 69, n. 1.

[2] The construction of such a curve has led us to evolve the concept of a practicable optimum which we have found useful in explaining the survival pattern of an industry. This will be at the scale, varying with the industry, where further economies of scale will be *relatively* negligible except for a very large increase in plant. In view of the flattening of the long-run technical cost curve in this neighbourhood, considerably smaller businesses could survive for relatively long periods if they adopted a relatively less generous obsolescence policy than the larger businesses.

ii. Normal Price-policy

A business man producing an article with a given specification will normally base his price on his costs of production.[1] He will be able to make fairly accurate estimates of his average direct costs, and, in order to get the quoted price, these will be grossed up by a definite amount, which following accounting usage we shall call the *average gross profit-margin* required per unit of product. The addition will equal the average contribution that the business man will require each unit of product to make towards covering the overhead costs of the business and making a profit. Whether the accounts for any period will actually show a profit or loss on balance will depend, of course, upon the extent to which the total sales[2] in that period have sufficed to cover the total overheads charged in the accounts as well as the direct costs.[3]

The exact method by which the average required gross profit-margin is calculated will vary between businesses, but the details do not affect the present analysis.[4] It is necessary only to take it as generally true that a definite addition will be made to current direct costs in order to get a price which will normally be adhered to (subject to relatively minor modifications or to short-term aberrations, one kind of which is discussed later). The amount added will formally include an allowance *as for* profit, although, as noted, the actual earning of a profit will depend upon the sales actually achieved.

Given, then, the business's costing rules and given the current level of its direct costs, the price which it will *normally* charge will be determinate, and will bear a definite relation to average direct costs. This principle of pricing will be called the *Normal Cost Principle*, since the business man, in fixing his price in this way, appears to act on the idea that such a price will normally enable him to cover his costs. The principle appears first to have been stated in economic literature in the article by Hall and Hitch, reporting some of the results of the inquiries of the pre-war Oxford Economists' Research Group.[5]

---

[1] Similarly if, as with some consumers' goods, he has to work to a conventional price, he will determine the specification of the product on the basis of his costs. It is for this reason that the length or weight of towels produced to be sold in some markets would provide a so-far unexplored index of the trade-cycle, since they have varied according to changes in the prices of materials or in wages, shortening in the boom and lengthening in the slump.

[2] We ought, strictly, to refer to output rather than sales because additions to stocks will also result in credits in the accounts. However, output decisions depend upon sales, and, since stock-policy is not considered here, it will be simpler to assume that stocks remain constant, when total output and sales will be equivalent.

[3] Given the price, the line showing total receipts will cross the curve of total costs at the point where sales would be just sufficient to break even with costs. This is the *break-even level of output.*

[4] The differences are, of course, important in practical work, and the reasons why a business should choose one particularly will be considered in my book.

[5] 'Price Theory and Business Behaviour', *Oxford Economic Papers*, Number 2, May 1939. The authors there named the principle as the *Full Cost Principle.* I have reluctantly

This costing-practice is stubbornly adhered to, as will be agreed by all who have discussed price-policy with manufacturers. Departures will certainly not be made for the sake of short-period gains but only in abnormal situations, and need not be regarded as weakening the normal cost principle as a general rule of behaviour.

Granted that the principle is so generally adhered to in manufacture, the mystery is that it has not yet emerged as a *postulate* of economic analysis. It should obviously be an accepted principle in pricing-theory.[1] Those economists who have thought about it have been too concerned with trying to explain it on the basis of the existing theory, instead of accepting it. Any explanation, of course, is wanted only for the purpose of constructing other and wider generalizations which will enable the principle to be incorporated more closely into the general framework of economic analysis. Scientific method would suggest that the right thing to do at the existing state of knowledge would be at least to accept the principle as a basis for further theory in its own field—the analysis of price-policy. An explanation is offered in the rest of this section, but it should be stressed that the validity of the principle as a theoretical postulate is independent of the argument.

As the Hall–Hitch paper has shown, business men offer at least the outlines of a rudimentary theory when questioned as to the reasons for their conduct. Economists have, perhaps, paid too little attention to the practical theorizing of business men—especially when produced in answer to the economists' own questions. Two things stand out in particular—the business man normally thinks long-run, and views the world he functions in as a *much* more competitive world than the modern economist is prepared to concede. In fact, a business man quite normally thinks that business is, in the long run, so competitive that there is a 'right' price, and that any serious departure from this will not be profitable in the long run, however attractive may be the gains of short-run price-changes.

In the long run, it is thought, a business which charges a higher price must have the ground cut from under its feet with the incursion of rivals into its market, and a lower price is thought to be unprofitable because the

---

changed the name, because the earlier term suggests, wrongly, that prices always tend to equal the full average costs of all business men. Even if they do so normally, they will not do so when output is running at a low level. The change in terminology has, at least, brought relief to one pedagogue.

[1] Economic literature had to wait for the inquiries of the Oxford Research Group and for the publication of the subsequent paper by Hall and Hitch for even a statement that it was the general practice. One would have thought that enough persons trained in economics had gone out into the business world for someone at least to have been impressed by the lack of any explicit marginal analysis and to have reported back from his frontier observation post that this is the custom, and that economic argument might at least consider adopting it as a principle of analysis. Further, it is now nearly a decade since the Hall–Hitch paper, and no attempt has been made to deny the general application of the principle which they then put forward, yet general economic analysis is still unaffected.

rightness of the price results from its representing normal costs. The business man therefore uses his costs as a guide from which to determine the right price in this sense, is sure that it will not pay him to charge more, and will only charge less when convinced by experience that only a lower price will keep him in his market, because he is less efficient than the rivals who are threatening him. Equally, while he will not normally cut his price *for the sake of* competition, he will be quite willing to offer a lower price than his rivals if he can justify it on his own costing rules, and will do so then because he will think himself able to sustain his attack on their position and to consolidate his own, his business thus being able to grow just because its lower costs will support it.

In more technical language, the business man is thus seen as thinking of his long-run demand curve as being infinitely elastic for a certain level of price, in the sense that any attempt to maintain a higher price will mean the loss of his market, since other producers will find it profitable in the long run to continue to supply the market at that price.

We must recognize that modern theory has led to erroneous views about the marketing situation confronting a typical manufacturing business. It is, of course, strictly true that no two businesses will produce exactly the same product. Even when the two products have identical specifications, they will be sold by different people, and all the personal circumstances surrounding the sale will be different. In fact, the multi-product business of the real world will usually specialize in some way, and any speciality will frequently be differentiated in specification.[1]

The fact of this separability of the markets of individual businesses appears to justify labelling the individual producer as a monopolist, rather than as a competitor, but the demand for the products of an individual producer is thought of far too much in terms of the short-run demand curves of the text-book consumers' market. Where the private consumer is concerned, it may not be false to regard his attachment as meaning that, in the short run, the individual seller will not lose all of his customers if he raises price. Equally, if he lowers price, he will not, in the short run, be able to attract all customers away from the rival sellers.[2]

---

[1] This is frequently the consequence of the way in which the founder set about getting *his* market and building up *his* goodwill, by doing anything which came his way in his line of business and, at the same time, looking round for any product on which he could specialize, in order to get buyers to give him more than the casual order and thus to become his 'customers'.

[2] Even in the consumers' market proper, however, the seller cannot rely upon the long-run support of his customers' preferences and habits. If he charges too high a price, rival sellers may well think it worth their while to incur the selling costs that would be necessary to transfer customer-attachments. Too high a short-run price thus means that his short-run demand curve will shift to the left in the long run, i.e. that the long-run demand curve will be elastic, its elasticity probably increasing with the height of his short-run price. This is one of the reasons why the Normal Cost Principle is found to be observed in retail

However, leaving the consumers' market on one side, the text-book view of preferences will not be true even for most makers of consumers' goods, let alone for the maker of producers' goods. The typical customer of a manufacturing business is another *business*, and, whilst such customers are no less subject to personal habits and preferences than the final consumer, they cannot, even in the short run, be as indifferent to the price which they are offered. To buy identically specified articles at a higher price than they could get them elsewhere would expose them in the long run to effective competition from those who buy more cheaply. Such business customers, moreover, are much better informed about costs and prices, and have a measure of any irrational preferences in the effects on their profit-and-loss account position. It is, accordingly, essential in the normal manufacturing business that its price should not be significantly higher than that which other businesses would quote for the same, or substantially similar, goods. Even where specialities are concerned (ruling out the case where they are protected by patent) they will normally be producible by other businesses—in fact, frequently are being produced, perhaps as non-specialities.

Apart from the businesses actually producing his type of article, a manufacturer will have two other potential competitors—new businesses, and new enterprise on the part of businesses who are producing other goods at the moment, but whose equipment and organization could be turned over to producing his type of product, if the gross profit-margin available to them were sufficiently attractive. The latter type of competition is important.

In an established business there will often be some drive from an existing department to expand by taking on new products which suit it. The mass of mature business leaders may be as conservative as are most men, but it is a characteristic of the business world that its rewards do attract the innovating mind, and in any industry there will usually be some up-and-coming young men with ideas who will force their way. In this way an efficient mature business is often subjected to expansionary pressure from its younger men. As one business man put it, looking back over the growth of his business, 'you have to expand to keep your young men'.[1] It should also be noted that many new ventures on the part of established businesses

---

pricing. It is realized that our general argument implies the abandonment of Chamberlin's 'large-group' analysis and the assertion of a kind of 'oligopoly' as normal. This will be defended in detail elsewhere.

[1] This paper has deliberately used the term business man in place of the term entrepreneur—it is a pity that some hardening of the linguistic arteries has removed the fundamental idea of enterprise from the traditional name. The characteristic of a business which is alive is the enterprise of its managers. The interest of the job is always tending to spill out in playing with new ideas and the idea of a new product quite naturally comes into the game.

have been caused by the desire to keep a key department more fully occupied in the face of a temporary depression in trade for the main product.

In economic literature a lot is made, and rightly, of the difficulty of new businesses getting sufficient capital and, hence, of the fact that established businesses may make large profits before new businesses start coming into the industry. When, however, the detailed history of individual businesses is examined, one becomes aware how normal it is for an established business to take on new products requiring broadly the same equipment, even if only in one department. In the course of time it is possible, in this way, for shifts in emphasis to lead to a business changing its typical product, and crossing the frontier from one industry to another. In fact, the frontiers of an industry are rarely fixed so firmly as they appear when, at any moment, one looks at the perhaps small number of apparently securely established businesses. Over a period one can often find a story of several attempts to expand into that product's market, some succeeding, some failing.

In short, the business man when he reports his world as competitive is generalizing from his experience and from his own attitude. A clue to the main characteristics of his approach, and of his innate competitiveness, is given at once when he is seen with the product of another business in his hand. It need not be at all the same product as he makes himself, so long as he has some acquaintance with the processes by which it is made. It appears almost an instinctive reaction for him, given the price at which the product is sold, to form a rough idea of the gross profit-margin which would be available to him, getting first an estimate of the cost of materials and then, if the net product seems attractive, of the direct labour cost. If the proposition continues to look attractive, there will follow at least the rough idea of the sort of additional equipment that he would need if he went for the new market. The process may well be left there in most cases, but there will remain the possibility of action being taken when circumstances are favourable, and a sufficiently attractive margin will always lead to action being taken somewhere. Of course, mistakes are made in the process, but innovation and the mistakes of innovation are part of the essence of business life.

In this way, the focusing point of competitive action is the gross profit-margin, and competition results, in the end, in stabilizing gross profit-margins in sections of industry which involve fundamentally the same sort of organization. So stable is it that, whilst one cannot make good guesses at the net profit-margin of a business, quite frequently after a walk round a works one finds oneself putting the business mentally in the same class as another which one knows, although that may produce a very different product and, on this basis, it may be possible to make surprisingly good guesses at the gross profit-margin.

It may help a little to understand the emergence of a business's costing rules if we consider the case of a brand-new product introduced exclusively by a particular business (it may, of course, be protected initially by patents). Theoretically, the new producer can charge what price he thinks best. He certainly will pay attention to short-run conditions of demand, especially if he is exploiting a patent valid only for a limited period of time. He may well set a very high gross profit-margin (but note that even his initial price-policy is reflected in his margin, which becomes the explicit part of his policy). There will be limits set to it by competition even here. Too high a margin will give an incentive for other business men to get round the patent, or even risk the chances of actions for infringement. The fountain-pen market appears to have known this sort of successful competition on occasion. If the product is not patented, or on the expiry of the patent, the product will be considered by other business men and some of them may come in, probably cutting away a bit of the original producers' market. To protect his position he will have to lower his price, i.e. granted that the specification of the article does not change he will have to reduce his gross profit-margin.

The competition of the new-comers and their zeal to get in, possibly counting on anticipated economies of scale to see them right when once they are in, will cause a fall in the gross margin. It may be set too low, in which case, sooner or later, some marginal businesses will be squeezed out, and it will be possible for the remaining producers to set a more remunerative margin. Granted the increase in their trade, and the reduction of costs through greater experience and technical development of the product, they may be able to get a higher margin whilst price is not raised. This has been apparently the normal story on the setting up of some of the modern British industries. The upshot is that, in the end, the gross margin of each producer will be set at such a level as preserves him from long-run competition in price by other producers, but which is at the right level for him to be able to continue to offer it in the long run against the others who might otherwise wish to come into his market.

Where the product is very standardized, and where technical conditions of production are similar for the producers, each of them will have very much the same rule relating his price to his direct costs, for price will not be able to differ appreciably for articles with identical specifications, and they will have very similar direct costs (there are usually minor differences in what are included in direct costs, resulting from differences in book-keeping practice rather than differences in principle). The more efficient firms will be charging a higher allowance for net profit, and at the given level of price they will tend to improve their financial position relatively to the other businesses.

In normal circumstances, price will be cut only when that is justified by reductions in direct costs. Once a normal cost rule has emerged in a business, price will be cut when these costs fall, for the producer will expect the lower price to continue to be profitable in the long run, and will regard any consequent cut in the prices of his rivals' products with equanimity. It is in this way that an efficient business emerges as the price leader in its market; the market will follow its price policy just because the price is lower than the competing firms would prefer to set on the basis of the normal costing rules and their higher direct costs. Where such price leadership has emerged, the leader has also been known to justify it on the grounds that a more generous margin would jeopardize the long-run position of the existing industry, in so far as it would provide a constant temptation to new entrants. Such a price leader will also try to prevent prices from being cut below the level given by its costing rules, but it will not be able to maintain that position very long if other businesses have in fact developed lower costs.

It may have more success against the price-cutting war of the kind described in the text-books, because it is generally recognized that this kind of price-cutting pays no one in the long run. It is recognized that any cut in price, if persisted in by any business, will lead to a general fall in prices—i.e. to a general cut in the gross profit-margins charged to get price. In the end, then, all producers merely get their share of the usually inelastic total market for the product, but the general level of price has fallen to a level which will not cover costs in the long run. Why, then, does such a situation occur?—usually because of the relatively desperate position of what is rightly called a weak seller. Given the prices at which a particular group of products is being sold, the general reduction in output which takes place in a slump will tend to bring at least some businesses to the point at which they are not covering their overheads. If their receipts do not cover the *paying-out* overheads such as overhead wages and salaries, their cash balances will be progressively depleted. It is then that a producer in a weak financial position, in danger of having to cut his organization or even to leave the market, while some at least of his rivals are in a stronger position, may try the remedy of cutting prices, knowing that there will probably be reprisals and that in the end he will lose his relatively improved share of the market, but hoping that the short-term increase in his receipts will enable him to hang out a little longer. In this way a depression, especially in industries with heavy overhead costs, may cause a departure from the normal costing rules. It may be noted that in such industries the costing rules themselves are usually based upon such a way of charging overheads as to permit of some relaxation, and thus to offer a deterrent against an abandonment of the rules.

Because such price-cutting does not pay, there are often attempts to prevent it by agreements, occasionally with financial sanctions. In view of the low elasticity of total demand, such agreements are not necessarily against social policy, provided that the gross profit-margin which they seek to enforce is only normal—and only in exceptional circumstances will it be possible to maintain an excessive gross profit-margin. It is to the community's interest that price should normally cover cost, and the normal pricing rules of business not only prevent price-cutting, they also prevent attempts to snatch temporary advantages from price rises in booms. If they prevent the share of profits from falling in the depression even more than it does, they also prevent it from rising so much as it would in a boom. A general departure from them in a depression would result in businesses cutting down their organizations, and dismissing overhead staffs when there is little prospect of alternative employment.

It is therefore suggested that the normal attitude to price agreements needs reconsidering. They are not necessarily anti-social, and it would be possible for social action independently to encourage the latent competition which generally exists. It would then not be possible for too high a profit-margin to be taken. In our experience it is difficult for price agreements alone to ensure that the participants retain excessive profits. Rings sooner or later bring into existence vigorous outsiders who grow and strengthen their position on the basis of finding it profitable to charge a lower price.

This is not to deny that what may be called monopolistic forces—the forces making it difficult to enter a particular market—will not affect the size of the margin which is taken as normal in the industry concerned, but it is possible greatly to exaggerate the size of the resulting monopoly net profits. In some at least of the industries usually referred to as quasi-monopolies (including cases where price agreements are known to operate), the figures for net profits do not appear to be large, bearing in mind the risks that necessarily attend industrial investment. British industrial history shows that there may well be some uncertainty about the security of investment in any industrial enterprise, however secure its position appears to be. The taking of only a reasonable gross profit-margin is one of the business man's ways of *reducing* that uncertainty.

Profits sometimes seem generous in the cases of well-established businesses which have been able to build up for themselves a rather specialized market, even in what may appear to be a competitive industry. Experience, however, suggests that such cases do not last, and that long-term forces do readjust the size of the margin. The tide of competition may leave little pools of abnormal profit behind it, but in the end they tend to disappear. In general, then, experience of industry does suggest that the business man is right when he sees his gross margin and his price as competitively determined. In our view, the newer developments in theory have caused economists to be too ready to regard manufacturing industry as a network of monopolies, and of deliberate restriction of output.

Certainly, on the basis of the theory which has been outlined here, the attempt to define as the degree of monopoly, the difference between marginal cost and price must fail. The horizontal nature of the direct-cost curves makes nonsense of any idea that in a purely competitive market (in Chamberlin's sense) equilibrium price would be that which equalled marginal prime costs—the standard for reference has therefore fallen to the ground. Equally it is necessary to attack the pricing rule which has been suggested for socialized industry—to make price equal to marginal costs would mean perpetual losses, and the nonsense of the rule would be obvious if all industry were socialized, for then how should we pay for the overhead factors in any industry?

These are wider topics, and a fuller treatment of them must be reserved for another occasion, some of them, in any case, calling for further research. It seemed, however, fitting at least to record the sense of competition which has emerged from the research which led to the theory which this paper has put forward. The firm conclusion for analysis is that the theory of normal behaviour should describe the business man as offering his product at a fixed price according to his cost rules, changing it only when his costs change. At that price the amount that he will sell will largely be determined by the National Income. The consequences for his profit-and-loss account will depend upon the amount that he sells, his total net profit increasing, or his net loss decreasing, with any increase in his sales. The gross margin that he takes in his price will depend upon competition, and hence upon the level of direct costs for the most efficient competitors. There is no reason to suppose that, in the case of the great majority of manufacturing businesses, competition will not keep that margin down to a reasonable magnitude. The theory of the gross profit margin should be one of the central parts of economic analysis, and should be formulated in terms of competition.

NUFFIELD COLLEGE,
OXFORD.

A REFORMULATION OF THE THEORY OF CONSUMER'S BEHAVIOUR[1]

By I. M. D. LITTLE

IN this article an attempt will be made to demonstrate, first, that the usual indifference-curve analysis of consumer's demand is logically unsatisfactory; secondly, that a theory of consumer's demand can be based solely on consistent behaviour; and, thirdly, that a synthesis of index-number theory, in its welfare aspect, and the theory of consumer's demand can be made.

Two criteria for being 'better off' are commonly employed by economists. The first, which may be called the index-number criterion, requires the fulfilment of two conditions—first, that $\sum p_1 q_1 \geqslant \sum p_1 q_0$ and, secondly, that $\sum p_0 q_0 < \sum p_0 q_1$. These two conditions together ensure that the individual in situation 1 could have bought the goods of situation 0, and that in situation 0 he could not have bought the goods of situation 1. The second criterion for being 'better off' is that the individual be on a higher indifference curve.

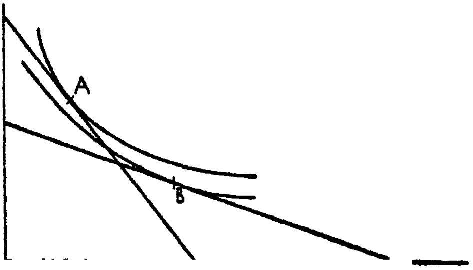

DIAGRAM I.

The question arises whether these two criteria come to the same thing or not. At first sight it appears not. It is possible to find a point *A* on a higher curve, and a point *B* on a lower curve, indicating two positions for which the index-number criterion does not hold, that is such that the individual in situation *A* could not have bought the goods indicated by the point *B*. Thus, in Diagram I, the individual in situation *A* could not buy the goods of situation *B*, and in situation *B* he could not buy the goods of situation *A*. By considering these two points only, the index-number criterion yields no relation of 'better than' or 'worse than' between them. Yet *A* is on a higher indifference curve than *B*, and, therefore, by the indifference-curve criterion, is better than *B*. It would therefore appear that 'on a higher indifference curve' is a test which may give results when the index-number test is silent. It will be demonstrated that this is not, in fact, true if one can take into account as many other points as one pleases. First, it may be possible, in such a case, to find another point *C* such that the index-number criterion shows that C could have been bought instead of A, and also that, if C had been bought in some different price-income situation, then B could have been bought instead. This, assuming consistent behaviour, would show that A is always chosen rather than C, and C rather than B. This follows because consistent behaviour, by definition, can be taken to mean: (*a*) if an individual once chooses A rather than B, then he will always do so, (*b*) choice is a transitive[1] relation, and (*c*) the individual never chooses a smaller collection when a larger[2] is available. It may be noted that the index-number criterion is not itself transitive, while the relation of choice is. It is because of this that a mediating point C can give us additional information.

---

[1] This reformulation is strongly suggested by the work of Professor Samuelson in *Foundations of Economic Analysis* (H.U.P. 1947). It rests on theorems stated by Professor Samuelson and, for this reason, all references are to Professor Samuelson's book. The logical differences are carefully noted. The indebtedness to Professor Samuelson is obvious. I am also greatly indebted to Mr. M. M. Crum of New College, Oxford, for his advice and criticism.

Now it may happen that A is on a higher indifference curve than B, but that no such mediating point C can be found which will allow the index-number criterion to tell the same story. Nevertheless, it may be the case that a series of such points can be found such that A is chosen rather than C, C rather than D,... and, finally, Z rather than B, and, therefore, A rather than B. Now we may go farther than this and propose that, if no such series of points can be found, then A is not on a higher indifference curve than B, by definition. The reversal of A and B in the criterion will similarly show, by definition, when B is not on a higher curve than A. Thus it is proposed that the indifference curve through a point B be defined as the boundary of the area marked out as 'better' than B by the repeated use of the criterion $\sum p_x q_x \geqslant \sum p_x q_b$ in the manner described above, where x is any point, and b is successively the original point and each mediating point. It would appear that it could also be defined, for the above purpose, as the boundary of the area marked out by the criterion $\sum p_b q_b \geqslant \sum p_b q_x$. It will be demonstrated later that these two boundaries coincide. It should here be noticed that these boundaries are not lines containing points between which the individual is indifferent. The concept of indifference is abandoned. The lines marking these boundaries will, therefore, be called behaviour lines.

This new definition of a behaviour line will be defended in three ways. First, the defects of the concepts 'preference' and 'indifference' will be discussed. Secondly, it will be shown that behaviour lines function in the same manner as indifference curves, and, thirdly, the logical simplifications which the change makes possible will be outlined.

Let us deal first with the concept 'preference'. The verb 'to prefer' can either mean 'to choose' or 'to like better', and these two senses are frequently confused in economic literature. The fact that an individual chooses *A* rather than *B* is far from being conclusive evidence that he likes *A* better. But whether he likes *A* better or not should be completely irrelevant to the theory of price, which can concern itself solely with market behaviour, and not with motives. In the theory of Consumer's Behaviour, as at present formulated, the preferences (in the subjective sense) which are discovered by asking questions are linked to market behaviour by the postulate that people do in fact try to maximize satisfaction. It is thus false to suppose that the theory, as it stands at present, is concerned only with choice. It is concerned also with likes and dislikes. However, it will be shown that it is possible to base the theory solely on actual choices, so long as perfectly consistent behaviour, in the sense defined above, is postulated. It may be noticed that this involves a slight stretching of the meaning of the verb 'to choose'. It will be said that *A* is chosen rather than *B* if, although *B* could have been bought, *A* is bought instead. This is in accordance with ordinary usage. It will also be said that *A* is chosen rather than *B* even when *B* could not have been bought, so long as there is some other collection of goods *C* which could have been bought, and which is always (that is, has been previously) chosen rather than *B*. The argument is similar if there are many mediating points like *C*. I think this slight stretching of ordinary usage is better than using the ambiguous word 'preference'.

---

[1] If ArB and BrC together entail ArC, then r is said to be a transitive relation.

[2] A collection of different kinds of goods is larger, if there is more of at least one kind, and no less of any kind. The condition that an individual chooses a larger collection of goods is not, strictly, a postulate. It is an analytic proposition following from the meaning of an economic good. A larger collection of things is not necessarily a larger collection of goods.

We come, now, to the concept of 'indifference'. The trouble with 'indifference', as with 'preference' in one sense, is that it is a subjective concept. There is certainly no obvious kind of market behaviour which can be called indifferent. How long must a person dither before he is pronounced indifferent? If a man dithers for ten minutes, and then chooses *A* rather than *B*, and as a result is said to be indifferent between them, we would have a contradiction between the two criteria for being 'better off'. The indifference-curve criterion pronounces him to be neither better nor worse off, while the index-number criterion maintains that he is better off than if he had to choose *B*. It is the subjective nature of the concept 'indifference' which prevents us from drawing up an indifference and preference map of objective preferences or choices.

We may now look at the normally prescribed method of drawing up such an indifference map. The individual is asked to put every possible pair of collections of goods into one of the three exclusive categories, '*A* preferred to *B*', '*B* preferred to *A*', or 'indifferent between *A* and *B*'. Therefore since the construction proceeds by questioning, it is logically possible that an individual for whom an indifference map had been constructed should proceed to behave in a manner which, although self-consistent, would be inconsistent with the map. Would economists in such a case have to declare that the man was irrational, and his behaviour inexplicable? The answer, as the theory is at present formulated, is 'Yes'. He would not be an 'economic man', because he would not be trying to reach the highest point of his preference field. But, surely, it must be said that there is something wrong with economics if it cannot explain perfectly consistent market behaviour?

We come now to the second part of our programme, which is to show that behaviour lines, as defined above, function in the same manner as indifference curves, and, therefore, that any theorem which can be deduced from the latter can be deduced from the former. This will be explained geometrically. The Diagram II is borrowed from Professor Samuelson.[1] The exposition differs in one important point, which will be explained.

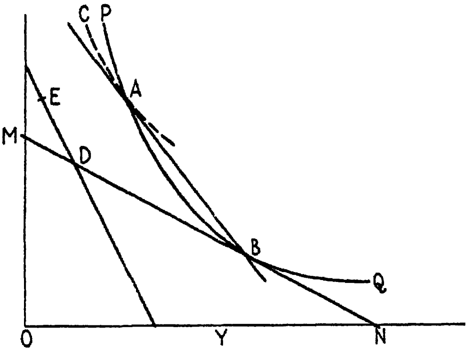

DIAGRAM II.

X and Y are two goods, quantities of which are measured along the two axes.

MN is a budget line.

B represents the collection of goods chosen from the possibilities indicated by all points in the area MON. It can therefore be said that B is chosen rather than any of these points.

Next, the budget line is rotated about the point B. The significance of this is that the price of Y in terms of X is altered, and at the same time the individual's income (in terms of X or Y) is altered so that he could still just, and only just, buy the same collection as before. In each of the resultant price-income situations a certain collection is chosen. The curve PBQ is the locus of the points representing these collections and may be called an offer curve. Take any one of these points A; the line AB is, then, the budget line on which A is the chosen point. A is therefore chosen rather than B, and similarly for any other point on the curve PBQ, and *a fortiori* for any chosen point above the curve, since more is chosen rather than less. The curve PBQ can only touch MN at B; everywhere else it must lie above it. If, as the budget line was rotated through B, a finite move to a point A was made only after the budget line had moved through some finite angle, then the points between A and B would never have been chosen, and the curve would not exist in the neighbourhood of B.[2] It follows that, if the curve exists at all in the neighbourhood of B, it is convex to MN at B. If we assume that all points would be chosen in some price-income situation, it can be shown that the offer curve must be continuous. Suppose that the individual remains at the point B as the budget line is rotated, until he makes a sudden jump to A. The offer curve would then be of the form $PACB$ in Diagram III. We know already that no point on the line AB, between A and B, is chosen when AB is the budget line. Any other budget line which intersects AB must cross the area chosen rather than B. But B is chosen rather than any point on the line AB between A and B. Therefore no such point can be consistently chosen. Any offer curve must therefore be continuous, on the assumption that all points would be chosen in some situation or other.

---

[1] Op. cit., p. 148.

[2] This would occur whenever an individual continued to consume exactly the same collection of goods in spite of a small change in their prices.

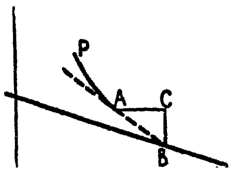

DIAGRAM III.

Now take the point A in Diagram II. In a similar manner, an offer curve can be drawn through A. Since it must be tangential to AB at A, it must cut the curve $PABQ$ and, therefore, enlarge the area within which it can be said that all points are chosen rather than B. This follows because any point C on or above the new offer curve is chosen rather than A (that is, the relation $\sum p_c q_c \geqslant \sum p_c q_a$ holds) and A is chosen rather than B (the relation $\sum p_a q_a \geqslant \sum p_a q_b$ holds). The point A can slide along the curve PBQ and generate a family of similar new offer curves. Similarly, points can be taken on these latter, and new offer curves generated, and so on. Eventually a limit will be reached, and the boundary of the resultant area within which it can be said that all points are chosen rather than B will, at present, be called the upper behaviour line of B.

Now let us take a point D on the budget line MN. We already know that B is chosen rather than D. Then rotate MN about D until D is the chosen point. We then know that D is chosen rather than any point E on the new budget line. Similarly D can, by moving along MN, generate a family of new budget lines, and so on as for the previous construction. Again, we will reach a limit which will be called the lower behaviour line of B.

It is evident that this lower line must be continuous, whether it is assumed that all points may be chosen or not. If it is so assumed, then, since *ex hypothesi* only one point is chosen on each budget line, this lower line must be concave; it must also be assumed to have a tangent at every point. Also, it cannot cut the upper curve anywhere, for in that case there would exist a point which was chosen rather than B, and rather than which B was chosen, and this contradicts the consistency postulate.

Before proceeding it should be noted that the difference between the above exposition and that of Professor Samuelson is that he states[1] that the position of points like A and C is determined by the condition of tangency with an indifference curve, and, presumably, implies that the slope of lines like *ED* is also so determined. In my exposition indifference curves are not mentioned. At this juncture Professor Samuelson states[1] that it is intuitively obvious that the two behaviour lines coincide and become the indifference curve through *B*.

---

[1] Op. cit., p. 149.

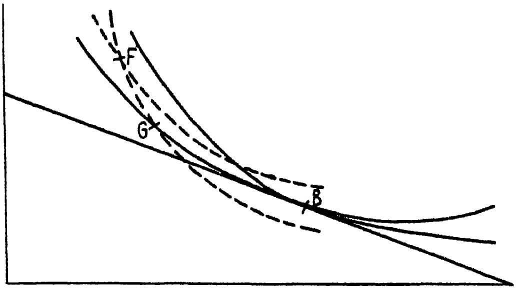

DIAGRAM IV.

This may be obvious if one has indifference curves already mapped out, but it does not seem quite so obvious if one has not. Although Professor Samuelson does indicate that an indifference curve may be traced out by the above method, the whole of his discussion implies that this would only be a tracing out of what is already known. It is, perhaps, significant that not only is Diagram II explained with reference to indifference curves, but also in his book it is included in the section on index-number theory and not in that on consumer's behaviour theory; also with reference to the latter theory, the maximizing postulate is held to be fundamental.[2]

It is required then to show that the two behaviour lines through *B* coincide. This will be shown by a *reductio ad absurdum.* In Diagram IV they are, therefore, drawn separate. The area between these two lines may be called *B*'s region of ignorance. Within this region take any point *F*, and construct behaviour lines of *F*. These latter are shown as dotted lines in the diagram. If we are to avoid contradiction, it is evident that the upper behaviour line of *F* cannot cut the lower behaviour line of *B*, for then *B* would be chosen rather than *F*, and vice versa. Inspection of Diagram IV shows that this is not an impossible condition. So far, there is no reason why lower lines and upper lines should not themselves intersect. Therefore, in Diagram V the lower behaviour lines of *F* and *B* are shown as intersecting at *G*. Consider the point of intersection *G*, and let us construct a lower behaviour line of *G*. Since both the lower behaviour lines of *F* and *B* will have different tangents at *G*, it is evident that the lower behaviour line of *G*, which itself has a tangent at that point, must cut one of the two other lines. In the diagram the dotted

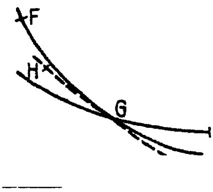

DIAGRAM V.

---

[1] Op. cit., p. 153.

[2] Op. cit., pp. 97 and 98. But in 'A note on the pure theory of Consumer's Behaviour', *Economica,* 1938, Samuelson takes a more austere attitude towards assuming the existence of indifference curves.

line is the lower behaviour line of G drawn intersecting that of B. It follows that B is chosen rather than G, and G rather than H. Therefore H must lie on or below the lower behaviour line of B, which is contradictory. Similarly, if the lower behaviour line of G was drawn intersecting that of F, a contradiction would arise. Therefore the lower behaviour lines of F and B must coincide, from which it follows that F is not in the region of ignorance of B.

A similar argument would be applicable to upper behaviour lines on the assumption that they are everywhere differentiable; but it is not required, since wherever F lies in B's region of ignorance the contradiction must arise as soon as its lower behaviour line is drawn. Therefore no such area can exist, and upper and lower behaviour lines must coincide.[1]

We have now reached the third part of our programme, which is to outline the logical simplifications which the above method permits.

To draw up an indifference map it was required that:

(1) The individual could place all pairs of goods in one of the three exclusive categories, '*A* preferred to *B*', '*B* preferred to *A*', or 'indifferent between *A* and *B*'.
(2) A larger collection be preferred to a smaller.
(3) The individual's preferences must remain constant.

Of these (1) disappears; (2) and (3) are reformulated in terms of choice.

---

[1] The above really completes the demonstration, but it may be added that in so far as the theory of consumer's behaviour is a scientific hypothesis which yields testable theories—and this is surely its proper status—the above result can be arrived at in a much more simple manner. Professor Samuelson states (op. cit., pp. 111, 115) that the following three testable theorems are all that are derivable from the theory of consumer's behaviour:

(*a*) demand functions are homogeneous and single-valued,

(*b*) $\dfrac{\partial x_i}{\partial p_i}+x_i\dfrac{\partial x_i}{\partial I} < 0$ (where I = income),

(*c*) $\left(\dfrac{\partial x_i}{\partial p_i}+x_i\dfrac{\partial x_i}{\partial I}\right)\left(\dfrac{\partial x_j}{\partial p_j}+x_j\dfrac{\partial x_j}{\partial I}\right)-\left(\dfrac{\partial x_i}{\partial p_j}+x_j\dfrac{\partial x_i}{\partial I}\right)^2 > 0.$

These theorems he also derives from the formula

$$\sum p\,\Delta x \leqslant 0 \quad \text{implies} \quad \sum (p+\Delta p)\,\Delta x < 0.$$

But this formula can be derived as follows from the postulates already suggested:

(1) given that an individual once bought the collection Q_1 rather than Q_0, i.e. given that $\sum p_1 q_1 \geqslant \sum p_1 q_0$, then it follows that if he ever buys Q_0 it cannot be the case that $\sum p_0 q_0 \geqslant \sum p_0 q_1$ because, in that event, he would have been able to buy Q_1, which, postulating consistent behaviour, would be a contradiction.

(2) Therefore, assuming consistent behaviour, and given that the individual has at some time bought Q_0 and at some other time Q_1, it follows that

$$\begin{aligned} &\textstyle\sum p_1 q_1 \geqslant \sum p_1 q_0 && \text{implies} && \textstyle\sum p_0 q_0 < \sum p_0 q_1 \\ \therefore\ &\textstyle\sum p_1 (q_0-q_1) \leqslant 0 && \text{implies} && \textstyle\sum p_0 (q_0-q_1) < 0 \\ \therefore\ &\textstyle\sum p\,\Delta q \leqslant 0 && \text{implies} && \textstyle\sum (p+\Delta p)\,\Delta q < 0. \end{aligned}$$

The above derivation is symbolically the same as that given by Professor Samuelson in his book (pp. 109–11), but the logic is altered to avoid the concepts 'preference scale' and 'expenditure minimization for a given level of satisfaction' which he uses, and which I find to be unnecessary. But in his article in *Economica*, 1938, the treatment is more nearly akin to the present formulation.

Further, to ensure consistency of behaviour with the preference field it was required that the individual always behave so as to reach the highest possible indifference curve. This requirement now disappears, because the highest point is now the point which was, in the same circumstances, chosen before. It is thus included under the consistency requirement. That maximizing individual, the economic man, who has caused much confusion, has at last disappeared.

The concavity of indifference curves also had to be postulated. But we have seen above that, at any chosen point B, the behaviour line in the neighbourhood of B must be tangential to the budget line. Therefore, if all points would be chosen in some market situation, it follows that behaviour lines must be everywhere concave. Their concavity thus becomes a theorem and ceases to be an axiom. With indifference curves, however, it was thought that concavity must be postulated in order to have a realistic map. If, for instance, the curves were anywhere straight lines, then in some price-income situation there would have been no determinate solution. In fact, there would have been actual indifference. But indifference, it was said, was never noticed; people always made definite choices. Therefore, indifference curves must be assumed to be everywhere concave. This argument was, however, a *non sequitur*, because there was no definition of indifferent behaviour, and so it never could be noticed. The result was that, in order to have an actual choice in all circumstances, it was, in effect, assumed that all possible collections could be chosen in some circumstances. On our formulation, of course, no question of indeterminacy can arise; the meaning of the concept 'choice' precludes it. But in order to prove *a priori* that, from observing market behaviour alone, it must be possible to arrange all choices in order on a *single* scale, it seems to be necessary to assume[1] that every point would be chosen in some price-income situation.

To sum up the logical advantages, the number of concepts is reduced by one and the number of axioms reduced from five to three.

It is also claimed that the new formulation is scientifically more respectable. If an individual's behaviour is consistent, then it must be possible to explain that behaviour without reference to anything other than behaviour. Someone, on the other hand, might object that market behaviour cannot be really explained by means of a map which is constructed out of nothing but that behaviour. The metaphor I have used to state this objection provides the answer. The terrain of England really is explained by a map of England. The map is constructed only by reference to this terrain.

---

[1] This assumption is one the reality of which may very well be doubted, and it would be practically impossible to test. In its absence, it is possible that there should be two points neither of which could be said to be chosen rather than the other. If all other points could be brought into relation with one of these two, then the individual would have two independent scales of preference. Ideally this difficulty could be got over by offering him, as a present, the choice of the two collections. In practice, this would be untrustworthy, because people take the exchange-value of presents into account. One would have to resort to asking the individual which he would choose if the collections could not be exchanged, i.e. which he prefers, in which case a maximizing postulate would become necessary. This theoretical difficulty must be admitted, but it has no practical importance.

The advantage is claimed for welfare economics that the two criteria mentioned at the beginning of this article for an increase in welfare reduce to one. Not only is it shown that these two criteria give the same answer, but they are shown to have exactly the same logical basis. Also, the index-number criterion is shown to be fundamental in the sense that the other is derivable from it. The empirical extension of the old indifference-curve criterion is precisely the same as that of the index-number criterion, if it is thought advisable to permit hypothetical choices in hypothetical price situations. This extension would require the additional postulate to the effect that the individual must behave in the way he says he would behave. Such an extension is not, perhaps, very useful when the only data we normally have are actual choices. Also, an individual may frequently not be able to say what he would choose in some hypothetical situation. Since the index-number criterion becomes fundamental, there is also the advantage of being able to explain in ordinary language, and with full precision, just what the economist's criterion for an increase in an individual's welfare is.

It may, on the other hand, be objected that a purely behaviouristic analysis takes all the welfare out of consumer's demand theory. But I do not think it was ever there. No one would to-day want to *identify* 'an increase of welfare' with 'a higher level of satisfaction as anticipated by the individual'. The transition to welfare economics can still be made in various ways depending on one's views. One could, for instance, say that a person is, on the whole, likely to be happier the more he is able to have what he would choose. Or, alternatively, one can say that it is a good thing that he should be able to have what he would choose.

It is tempting, in conclusion, to make some evaluation of the usefulness of the theory of consumer's behaviour. First, one may well ask whether the economic behaviour of an individual is likely to be consistent. Now, in order to apply the theory, one must consider some period of time within which the collection of goods is supposed to be bought. If the period is very short, some goods which are only bought occasionally would not be included; also, if the period is short, the person's tastes will be influenced by what he has consumed in the preceding period; in other words, he may have a hangover, and therefore display inconsistency. If, on the other hand, the period is very long, the person's tastes may suffer a change while the collection is being bought, in which case there will be inconsistency between this period and the next. So considerable doubt can be thrown on the reality of our assumptions. Secondly, the theorems which can be deduced are not very important or interesting; nor would they be at all easy to test in practice. It may be asked, therefore, why the theory of consumer's demand has been considered to be fundamental, or even important. It cannot be for empirical reasons, because very little follows about individual demand functions, or about the market demand functions, in which the practical economist is interested. It also cannot be for logical reasons. There is absolutely no logical reason why market demand curves should not be considered to be the basis of price theory.

Now, whereas there appears to be no reason why price theory should concern itself with the individual, it is certainly true that most people would consider it essential to get back to the individual for the purpose of welfare theory. Formerly, price theory and welfare theory were interwoven. The individual was assumed to be always striving to achieve the greatest amount of satisfaction possible. So long then as individuals were wise and omniscient, their free choices, which helped to determine prices, resulted in their obtaining the maximum amount of satisfaction possible in the given circumstances. It is thus easy to see why it was then considered that the theory of consumer's behaviour was of fundamental importance. Welfare theory and price theory were partly disentangled by the development of the indifference-curve analysis; but not entirely, because 'on a higher indifference curve' has remained a criterion of being 'better off'. It is suggested that the present reformulation completes the disentanglement, and shows that a subjective interpretation of the theory of consumer's behaviour cannot be held to be of importance for welfare reasons.

Another reason why it may be believed that the theory of consumer's behaviour is, in some sense, fundamental is that there is a tendency to think that one has not arrived at a proper explanation of 'macroscopic' events if they cannot be shown to be determined by 'microscopic' events about which general laws can be formulated. This tendency probably results from the success of these methods in some branches of physics. But it is only useful thus to push back the analysis if laws can be formulated about the behaviour of the 'atoms'. The theory of consumer's behaviour is, however, a deductive system based on a postulated consistency, and, otherwise, it contains no hypotheses based on induction which would enable us to predict the behaviour of prices. The behaviour of human beings seems, in fact, to be less predictable than the behaviour of aggregates of human beings, and, to the extent to which this is true, the theory of consumer's behaviour must, at least as far as positive economics is concerned, continue to be nothing other than a purely logical exercise, because, in price theory, it is the aggregates, and not the individuals, which are of interest.

NUFFIELD COLLEGE, OXFORD.

THE NATURE OF INTEREST-RATES

By G. L. S. SHACKLE

Prefatory Note. The question what is the *nature* of interest, what psychological, institutional, or technological realities it manifests and corresponds to, is surely the first one we must answer in seeking to know how interest-rates are determined. The time-preference hypothesis had all the incisive formal brilliance characteristic of marginal value analysis in general, but it belonged to that world of thought which virtually ignored *uncertainty*: which assumed that men choose between alternative trains of consequences each known in its entirety through all relevant future time. The liquidity-preference hypothesis (to the authorship of which Lord Keynes in the *Treatise* and the *General Theory*, and Mr. Hicks in his intervening *Suggestion for Simplifying the Theory of Money*, have, I suppose, independent and not very disparate claims) utterly transformed the matter (Professor Robertson has pointed out to me Lavington's claim in this connexion). Now M. Pierre Massé in an article which, I think, deserves very high praise: 'La Théorie générale de J. M. Keynes et le problème de l'intérêt' (*Revue d'Économie Politique*, janv.–fév. 1948), seems to wish to synthesize the two views. I differ from him on two counts: first, the language of probability theory seems to me wholly unsuited to the analysis of uncertainty. (This I have tried to show in a book not yet published.) But that is not the point here. M. Massé's contention that the propensity to save has, after all, a direct bearing on the interest-rate impelled me to try to re-think the theory of the nature and determination of interest-rates from fundamental considerations. This article is an attempt to express the main outline of the result. Full development and the pursuit of all the questions this topic raises would occupy a whole book.

I wish to express my warm gratitude to Messrs. Pember and Boyle, of Princes Street, E.C., for giving me access to their records of security prices and yields, and especially to Mr. F. R. Althaus of that firm, to whose most kindly helpfulness and illuminating comment I am greatly indebted.

CONCERNING a stock of wealth, there are two kinds of decisions which its administrator must constantly be taking:

First, in what proportions it shall be composed of various things.

Second, how fast he shall augment or diminish it.

I shall describe some simple assumptions and my grounds for believing them true enough to be interesting and useful; and then show what formal criteria, on those assumptions, are logically appropriate to the decisions. Having arrived at the principles on which the individual decision-maker[1] will administer his wealth, I shall consider what will happen when the many simultaneous individual decisions are acted upon or attempted. The conclusions thus reached will necessarily include a theory of the determination of the system of interest-rates.

My basic assumptions in brief are that an individual likes to possess things, the more the better; and likes to consume things, the more, in a given time, the better; and that he cannot both eat his cake and have it

---

[1] In the notion 'individual decision-maker' I include one who directs an enterprise as well as one merely concerned with his personal fortune.

(but I think this notion needs special analysis to make it precise, and this I have attempted). Also that an individual's judgement as to how much wealth is represented by a given collection of things depends on his system of expectations, a system in which different hypotheses are given different degrees of influence on his decisions according both to the content and to the plausibility of each; and that it is the amount of his wealth thus reckoned that he will seek to maximize in choosing at each instant the proportionate composition of the collection; that hopes and fears thus play a central role in his personal relative valuation of different assets, and therefore in determining exactly what degree and direction of pressure on the market where the interest-rate is determined is contributed at any instant by each individual.

Why should a man accumulate or keep in being a stock of wealth? The feelings afforded by the possession of such a stock are essentially, I think, the feelings of security, power, and freedom; in a word, the feeling of *independence*. Those three things are in a large measure all aspects of the same thing, for full freedom requires a degree of security from the consequences of other people's decisions, and also some power to execute one's own ideas. All this it is convenient to label with the one term 'possessor-satisfaction'. Possessor-satisfaction is to be thought of as an instantaneous *intensity* which will be greater in that one of two otherwise identical personal situations in which the person's stock of wealth, by his own reckoning, is the greater.

A stock of wealth can consist of concrete objects (or systems of objects), or of abstract, generalized claims. Of the latter there are two kinds, dated and undated. *Undated* claims include banknotes and coin, credit balances at the bank, and unused permission to overdraw. All these I shall comprehend in the term money, but instead of speaking of a 'quantity of money' I shall often speak for the sake of vividness of a 'number of banknotes'. *Dated* claims are promises to pay stated numbers of banknotes at stated dates. In this first, theoretical, part of this paper I shall refer to all such promises as 'bonds'. All *concrete* objects which embody wealth I shall call 'equipment'.

A stock of wealth, then, consists of banknotes, bonds, and equipment. An object belonging to any of these three classes of objects I shall call an asset. The quantity of equipment comprised in any particular stock can be positive or zero, the quantity of bonds and that of banknotes can be positive, zero, or negative. Since the market valuation (in banknotes) of any asset is merely a weighted average of personal valuations, any decision-maker's personal valuation of any asset except a banknote can differ from the market valuation. Finding himself at any instant in possession of a *given* collection of assets, a decision-maker will exchange some of his assets for others so as to get that collection of assets to which he assigns the highest *personal* valuation amongst all those collections whose *market* valuation is equal to that of his given collection. For thus he will maximize his possessor-satisfaction.

Ready money, banknotes in hand or available at the bank, is in ordinary circumstances the most reassuring form of wealth. Its power to discharge debts and contracts stated in terms of money is secure against market-value changes. Its equivalent in real goods (i.e. in a well-assorted bundle of many varieties of real goods) is felt to be more dependably stable than that of systems of concrete, specialized property, or shares in them. It is available without delay. The man who has decided to keep in hand such-and-such a number of banknotes knows that, until he changes that decision, he will have that number of banknotes at his command for use at any instant, without need to know in advance when that instant will be. A bond is a claim to receive stated quantities of banknotes at stated future dates. If therefore deferment was of no account, a bond (given perfect trust in the honesty and solvency of the debtor) would be valued at just the sum of these quantities of banknotes. When there is a free market in bonds, the creditor can escape from the fixity of these dates of receiving payment, and can look on the bond as a source of banknotes available, by sale of the bond, whenever he may need them: but only at the cost of not knowing in advance how many banknotes he will receive. A man who accepts a bond in exchange for banknotes is therefore exchanging a *known* for an *unknown* quantity of banknotes. (For let us remember that by *banknotes* we mean money *in readiness*. When we ask, concerning a bond: What does this bond represent in money available instantly at an *unforeseen* instant, as banknotes themselves are available? the answer is: an *unknown* number of banknotes.) He will not make this exchange unless he has grounds for hoping that the unknown quantity will prove to be larger than the known; and amongst these grounds he will virtually always require (when considering whether or not to buy an existing bond or to allow a borrower to create a new one) that the quantities of banknotes due from the debtor shall add up to a total larger, by an amount depending on circumstances which we must analyse, than the amount he pays for the bond. The appropriate measure of this premium is the *yield*, defined as follows: Suppose that Q_k is the number of banknotes due from the debtor at a date of k time-units' futurity, such payments being due at the dates for which $k = 1, 2, \ldots,$ up to L. Then if P is the number of banknotes paid at the present instant for the bond, and if r is such that

$$P = \frac{Q_1}{1+r} + \frac{Q_2}{(1+r)^2} + \cdots + \frac{Q_L}{(1+r)^L}, \quad \text{or} \quad P = \sum_{k=1}^{L} \frac{Q_k}{(1+r)^k},$$

r per time-unit is the yield of the bond at the present instant. The payments Q may be all equal and may stretch indefinitely into the future, being terminable only by the debtor buying back the bond at whatever may then be its market price. Or there may be a terminating series of equal payments with a final larger payment representing 'repayment of principal'. The equal payments in either case are conventionally known as 'interest payments'. But there is no reason in theory why the payments should not be all unequal, and even irregularly spaced in time, provided the dates and amounts are specified at the outset: the yield as defined above still measures the compensation offered to the buyer of the bond for accepting its terms. In actuality there are at all times on the market fixed-interest securities offering a great variety of time-schedules of payments. Some are due for completion of payments by the debtor within a few months (treasury or other bills) or a few years (redeemable government securities which are approaching the latest date for 'repayment of principal'); and the length of time over which future payments will stretch ranges up to perhaps forty years for the 'longest' redeemable gilt-edged. Lastly, there are the 'irredeemables', the holders of which are promised only an indefinite series of equal 'interest' payments until the securities shall be bought in by the Treasury at an unspecified date. In the rest of the formally theoretical part of this paper I shall assume, for the sake of simplicity and vivid presentation of main ideas, that only *one type* of fixed-interest securities exists, viz. irredeemable securities issued either by the Government or by individual enterprises. I shall suppose that these are issued in equal nominal units each called a 'bond'.

The hazard accepted by the buyer of a bond is that its market price in banknotes may fall after he has bought it. Even if he should not be compelled to sell the bond while its market price is low, still the fall must in some sense be regarded as the lender's loss, since if he had waited he could now have had the bond, as in fact he has, and *also* that number of banknotes by which its price has fallen; but these he has lost. This hazard has of course its counterpart in that the price of the bond may rise. It is the net effect of these two hazards, on the mind of a potential holder of the bond, which must be compensated by the yield.

Let us summarize our comparison of the characters of bonds and of banknotes. *Dated* claims, just like undated money, are abstract, generalized purchasing power, but they do not, as undated money does, combine *certainty* of amount with *freedom* of date of spending. If the holder of dated claims could feel *certain* that he would need banknotes in just those numbers at just those dates when they are due from the debtor (or at any rate, no sooner), those dated claims would represent a *known* number of banknotes. (But he cannot be justified in feeling certain of this.) Under all other circumstances dated claims represent an *unknown* number of banknotes. For this hazard the compensation, which induces people to hold some dated money, is pure interest as measured by the yield.[1]

A piece of concrete equipment resembles a bond in being valued for the sake of instalments of money to be received at dates in the future. The difference is that with concrete equipment the sizes of the instalments which will accrue at particular dates are not trusted promises but mere conjectures. This set of conjectures can be described, and a value in terms of banknotes in hand assigned to it, in various ways. For example, a single number can be assigned to each future date as the number of banknotes expected to be earned by the equipment in the unit interval ending at that date, and this number can then be multiplied by a coefficient to allow for the doubt in the owner's or potential owner's mind concerning the suitability of that number. The result will be a set of numbers (one for each relevant future date) each of which can be treated as though that number of banknotes were *certain* to be received at the date to which it belongs, and the piece of equipment will then be valued, by the individual in whose mind this process of conjecture and allowance for doubt has gone on, as though it were a bond promising those numbers of banknotes.

A collection of given quantities of banknotes, bonds, and various systems of equipment (or shares in their equity) will have at any instant a market value and also a personal valuation assigned by the individual administrator of this collection, both these valuations being expressed as a number of banknotes. This personal valuation must not only take account of the administrator's system of expectations, in which rival hypotheses count for more or less with him according both to the content and to the plausibility (in his own view) of each, but must also take account of his own attitude to this state of uncertainty, the gratification or distress it causes him. Now it would be possible for him to turn the whole collection into banknotes at its market valuation, and there would then, of course, be no difference between personal and market value. To depart somewhat from this state of affairs by parting with some banknotes in exchange for bonds may give him a collection with a higher personal valuation. It may even be that, amongst all the different collections containing only banknotes and bonds, which he could obtain at the current market price of bonds, the one consisting wholly of bonds will have the highest personal valuation. It is evident, however, that the price of bonds cannot remain at a level which makes this true of *all* individuals.

---

[1] Compensation for any doubt which there may be concerning the honesty or solvency of the debtor I do not regard as pure interest. Most of the 'interest' exacted by a money-lender is not 'pure interest' in my sense.

Except in so far as bonds are bought from the banking system, the banknotes paid for them remain in the possession of some individuals (those who sold the bonds) and the price of bonds must accordingly rise to a level where some individuals cannot increase their personal valuations of their own collections of assets by parting with banknotes in exchange for bonds, but will be content to retain the banknotes. Can we explain on reasonable grounds why there should always be some level (varying of course with circumstances) at which the price of bonds will satisfy this condition? The higher the market value of a bond, the less plausible (other circumstances given) is the hypothesis that the next movement of that price will be a further appreciable rise. But only an individual with whom this hypothesis carries some weight will value bonds higher than their immediate market price. Must we then suppose that some individuals will at any instant hold nothing but bonds, others nothing but banknotes? No. For let us look at the matter thus.[1]

We have supposed that any uncertain outcome, such as the immediately future behaviour of the price of bonds, is represented in an individual's mind by a set of rival hypotheses to each of which he attaches some greater or lesser degree of importance according both to its content[2] and to the plausibility he concedes to it. Let us suppose, then, that embracing in one glance all the hypotheses of *loss* from the holding of a single bond, through a fall in the market value of bonds greater than is made up for by future interest payments, he can select some one amount of loss (say, L pounds) which sums up for him the whole of this part of his system of expectations; and let us call this amount of loss his 'loss-epitome' for one bond; and that, similarly, he can find some one amount of gain (say, R pounds) which represents for him the whole of the other part of his system of expectations, including any payments due from the debtor within the time interval over which he looks ahead; and let us call this his 'gain-epitome' for one bond. And in a Cartesian diagram[3] let the abscissa measure losses and the ordinate measure gains, so that the loss and gain potentialities of a single bond, as seen by our individual wealth administrator in the circumstances of some one historical instant, will be represented by a point with abscissa L and ordinate R. The loss-epitome of a holding of N bonds will evidently be an abscissa NL, and its gain-epitome an ordinate *NR*. Now every point in the plane of our diagram[1] will represent some conceivable combination of loss-epitome and gain-epitome, a combination which could become actual in appropriate circumstances for some particular individual with some particular holding of bonds. Selecting any one such point, we can discover all those other points which represent situations neither more nor less agreeable to the individual, and join all these points by an *indifference curve*, which, unlike the indifference curves of value theory, will evidently have a positive slope. All points in the relevant quadrant of the plane will lie on one or other such curve, and if curve *A* (say) lies *above and to the left* of curve *B*, every point on curve *A* will represent a situation preferred to every situation represented by a point on curve *B*.

---

[1] The analysis which here follows is an outline sketch, adapted to the very small space which can be spared for it in this article, of a much more fully worked version first presented in my article 'An Analysis of Speculative Choice' (*Economica*, February 1945) and in a further improved form in my book *Expectation in Economics* which has been accepted for publication by the Cambridge University Press. For the sake of extreme brevity I have here laid aside the precisely defined concepts of my full technique and have adopted fresh terms for the roughly indicated ideas which take the place of the exact ones.

[2] The proposition that it contains, without regard to the plausibility, in the given circumstances, of this proposition.

[3] See the accompanying Chart 1.

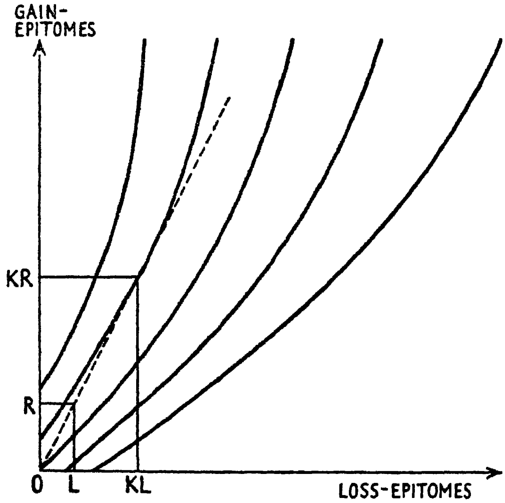

CHART 1

The indifference curves will evidently slope up from left to right, but what will be their shape? A loss which is large in relation to his total wealth may appear to the individual a disastrous contingency, the fear of which can by no means be neutralized by an equally vivid hope of an equally large gain. But when contingent loss and contingent gain are both small in relation to his total wealth, equality between them may represent a situation neither preferred to nor rejected in favour of *zero* loss and gain, this latter being the situation secured by a decision to hold *banknotes only* and represented by the *origin* of the diagram. Thus it seems plausible that the particular indifference curve which passes through the origin will have an *increasing* slope, equal to unity near the origin and becoming steadily steeper; and that the shapes of all the other indifference curves will somewhat resemble it.

---

[1] That is, in its relevant (upper right-hand) quadrant.

In our diagram the point (L, R) represents the subjective uncertainty-situation at a particular historical instant of an individual who then possesses *one* bond, perhaps some banknotes, but no equipment. This point (L, R) will be the same no matter how many or few banknotes he possesses in addition to his one bond; for the holding of banknotes cannot give rise to either loss or gain in terms of banknotes. The point with abscissa NL and ordinate NR will represent the situation he secures by holding N bonds,[1] and again it will be unaffected by the number of banknotes he also possesses. Thus if, expressing the market value of his stock of wealth (whatever its actual composition) as so many banknotes, the individual imagines himself to carry through a process of successive exchanges of banknotes for bonds, this process will carry his situation along a *straight line* of slope R/L passing through the origin. If this line has a slope greater than the slope, at the origin, of the indifference curve which passes through the origin, and the slopes of all indifference curves increase from left to right, then the straight line will have somewhere a point of tangency with an indifference curve. That point (if his wealth is enough to carry him so far along the straight line) will represent the situation preferred by the individual to all others which he can attain, and will determine the preferred composition of his stock of wealth: so many bonds, so many banknotes.[2] The amount of his wealth (at a given market price of bonds) will determine how far along the straight line he can go: if he cannot reach a point of tangency with an indifference curve, he will have to be content with the highest curve encountered by that segment (of the straight line) over which he is free to move.

In the preceding analysis we have assumed that in deciding on the composition of his stock of wealth the administrator limits himself to a choice amongst combinations in various proportions of banknotes and bonds, and excludes from it equipment or equities. (Though we did this for analytical convenience, such a limitation is often compulsory for a trustee.) If we now relax this assumption, there are two alternative ways in which we can suppose the administrator to choose the composition of his stock of wealth when equipment of any number of different kinds, each signifying for him particular possibilities of gain and loss, can be included in it. We could suppose him to use a mental process equivalent to an elaborated version of the gain-and-loss indifference diagram described above. Instead of an opportunity-curve consisting of a single straight-line segment whose end-points represent respectively the holding of his whole fortune in banknotes or in bonds, and whose intermediate points each represent some combination of banknotes and bonds, we have then in addition a swarm of points each representing the holding of his whole fortune in the form of shares in some concrete enterprise. On the diagram a point can then be found corresponding to *any* proportionate combination of any number of different assets, and the preferred combination can be determined as before. This analysis I have worked out in detail elsewhere.[1] But for our present purpose another treatment is more directly illuminating. We suppose the individual to divide future time into unit intervals and to assign to the end-date of each such interval a single number to represent his 'best guess' as to the net[2] earnings of any enterprise or piece of equipment in that interval, and then to multiply each such number by a coefficient to allow for his doubt of the appropriateness of that number. The coefficient will take account not only of his uncertainty itself regarding the size of the instalment, but also of his attitude to that uncertainty, and the result will be a series of numbers, one for each relevant future date, which series will be on the same footing as the payments promised by a bond. This series of dated, uncertainty-adjusted instalments, and the market-price of the equipment, will together make possible the calculation of an exact analogue of the *yield* of a bond. This analogue or something closely resembling it has received from various writers a number of different names, such as 'rate of return over cost' (Irving Fisher) and 'marginal efficiency of capital' (Lord Keynes). Not all of these writers have sufficiently emphasized the need to allow for the individual's uncertainty and his attitude to uncertainty before calculating this measure, and before stating the proposition that equipment of each different kind will tend to be produced at a time-rate such that, in the given conditions of its production prevailing at any instant, the cost per unit will be pushed up to such a level as to make the marginal efficiency equal to the yield of bonds. It will be evident, I think, that the marginal efficiency of a piece or system of equipment is subjective and perhaps will not be set at the same figure by any two individuals who form an opinion on it. The marginal efficiency referred to in the above-mentioned proposition must be that of the *most sanguine* individual whose circumstances, in particular his command of wealth and his *expertise* in the line of production concerned, imply that he is potentially in the market for such equipment.

---

[1] N being thought of as a variable, no particular point (NL, NR) is marked on Chart 1.

[2] In Chart 1 the point of tangency is shown as corresponding to a holding of K banknotes.

In so far as the purpose of the coefficients of allowance for doubt is to allow for the individual's *liking or aversion for uncertainty*, the scaling-down effected by them may, it seems plain, have to be comparatively severe when a large fraction of his total stock of wealth is embodied in one kind of equipment; for then his position may be more vulnerable to hazards whose importance *per unit* of equipment is *given*. Thus for a given individual with given expectations there is likely to be some relation, between the quantity he possesses of any particular kind of equipment and the size of his total stock of wealth, which will make the marginal efficiency of that equipment equal to the yield of bonds. In order to adjust his actual situation to meet that condition, he may need to increase or decrease his holding of that kind of equipment. Moreover, he may even need to hold a *negative* stock of bonds: to become a borrower. This latter is the position sought by an enterprise when it issues debentures in order to hold a larger quantity of equipment than can be bought out of the equity capital alone. But the larger the ratio of the enterprise's debt to its assets, the larger will be the ratio of the periodic payments due on that debt to the earnings that can reasonably be expected from its real equipment, and the smaller the decline in its rate of profit which will compel it to default on those payments.[1] Thus there is plainly a point beyond which the buying of further equipment with borrowed money will reduce the administrator's subjective marginal efficiency judgement of this equipment below the interest-rate (yield) at which he can borrow. This idea can be very conveniently shown on the uncertainty indifference-map. The point representing the holding of a negative stock of bonds (the situation of an enterpriser who has become a net borrower in order to hold more equipment) can be constructed from the point which would represent the holding of the same equipment with a *zero* stock of bonds (i.e. without borrowing) by increasing the loss-epitome, and decreasing the gain-epitome, of this latter situation by the amount which the borrower is bound to repay (making due discounting allowance for the date of repayment).

---

[1] In 'An Analysis of Speculative Choice', *Economica*, February 1945, and in an improved and developed version, Chapter IV of my manuscript *Expectation in Economics*.

[2] By *net* earnings I mean sales proceeds of product less expenses of operating the equipment.

I turn now to the second kind of decision which the wealth administrator must take: how fast to augment or diminish his stock of wealth. Here we wish to know two things: first, on what formal principles, and by reference (in applying those principles) to what empirical facts will he decide; and second, whether and by what mechanism the simultaneous decisions in this matter of all individual wealth administrators considered together will affect the interest-rate (the system of rates).

May we not say that the *possession* of a stock of wealth gives its possessor (through whatever psychic mechanism) some experience whose intensity would be somewhat increased by the addition of one banknote's worth to the stock, and increased again, but by a slightly smaller amount, by the addition of a second banknote's worth, and so on? Let us call this experience *possessor-satisfaction*. But what do we mean here by possession? We do not merely mean 'having possessed *until* the present instant'. For possessor-satisfaction to be yielded there must also be the intention that possession shall continue for some time into the future. For how long? The word *decision* plainly covers a considerable range of meaning, from mere intellectual recognition of the superiority of some one line of action over others, to an actual mustering of moral, mental, and bodily forces for the beginning of its execution. But properly used it must always, I think, include the idea of *mentally committing oneself*, so that the *immediate* stage of the line of action decided on is treated and felt as an accomplished act. For each particular person in any given field of action this 'immediate' stage will have, more or less consciously, some particular extension from the 'present instant' (when the decision is taken) into future time. Let us call this time-interval the *decision-interval*. It seems to me reasonable to say that the decision-interval is the least interval for which possession must be intended to last in order that it may generate possessor-satisfaction, and also the longest interval for which 'intention' means something that can generate such a feeling. The decision-interval is thus a means of dividing a stock of wealth into two portions: that portion from which the owner can, at the given historical instant, derive possessor-satisfaction and that portion from which he cannot derive possessor-satisfaction because he has decided to delete it from his stock within the decision-interval. He will, of course, require any assets thus dispensed with to yield satisfaction in another way: they will be exchanged for 'food and drink', which he can consume and thus derive *consumer-satisfaction*.[1]

---

[1] This is Mr. Kalecki's *principle of increasing risk*, which I should formulate by saying that the pace of growth of an enterprise's total assets must not too greatly exceed the pace of ploughing back of undistributed profits or raising of *equity* capital.

If the wealth administrator believes that during the decision-interval some number N of banknotes will accrue to him as income, he may decide to spend a portion $(N-A)$ of these during the decision-interval on 'food and drink' and add the remainder, A, to his stock of wealth. Then by definition of the interval, this addition will be looked upon by him as something accomplished, which will therefore augment his intensity of possessor-satisfaction. In effect, we need make no theoretical distinction between the cases where N is zero, so that any spending on consumption must actually reduce the stock of wealth below its level at the instant of decision, and the case where N is positive. The wealth administrator, according to this view, will have in mind the size that his stock of wealth will have at the *end* of the decision-interval. Any consumption he decides to make within the interval will reduce that size below what it could otherwise have been. If this reduction is by B banknotes, we may reasonably assume that the greater is B, the more intense will be his *consumer*-satisfaction and the greater also will be his corresponding *loss* of intensity of *possessor*-satisfaction; but that the greater B is, the smaller will be the *increment* in his intensity of consumer-satisfaction, and the larger will be the *decrement* of his intensity of possessor-satisfaction, resulting from a change of decision from B to $(B+1)$. Then his combined intensity of possessor- and consumer-satisfaction will be maximized if he so chooses B that a change of decision from B to $(B+1)$ will increase his consumer-satisfaction by no more nor no less than it reduces his possessor-satisfaction.

---

[1] I use the expression 'food and drink', here and hereafter, as a more vivid substitute for 'consumption-goods'. I hope that such avoidance of words whose impact on our minds has been dulled by continual and various use may make the argument more readable.

We have answered the question: On what formal principles will the individual decide how fast to augment or diminish his stock of wealth? We can now pass to the study of a *system* comprising a number of such individual decision-makers, and having also a banking system, but no taxation. There is a formidable difficulty. Hitherto we have been concerned with the *thoughts* occurring in one decision-maker's mind *at an instant* or, perhaps we should say, in a very short interval whose length can be neglected. We have spoken as though in reaching his decisions he *knew* how many banknotes would accrue to him as income in the decision-interval, and at what price he could sell or buy a given quantity of bonds. In a realistic model we cannot assume these things known to him when he takes his decisions, for they depend on the *simultaneous* decisions of other people. Nevertheless, something may perhaps be learnt from a model in which the individual does in effect base his decisions on known income and known terms of borrowing and lending (i.e. known terms of sale and purchase of bonds). Let us suppose that at the beginning of each 'week' every decision-maker (whether on his own behalf or as head of an enterprise) must send in to a central office a schedule in the form of a square array, each column of which is headed with a number representing a hypothesis as to the aggregate income of all decision-makers in the system for that week (the hypotheses covering a suitable range from low to high), and each row of which is labelled with a number representing an hypothesis as to the yield of bonds in that week. Any number inserted in a row and in a column of this array will stand for the individual's conditional promise to spend during the week, either on 'food and drink' or else on equipment for a net enlargement or improvement of his productive facilities, or on some combination of such consumption and net investment, that number of banknotes. The conditions for each such promise are that the aggregate income and the yield of bonds should be those named at the head of the column and the end of the row in which the number stands. The sum of the numbers inserted in any one place in the array by all the individuals who together compose the system is then the sum of their conditional promises to make so-and-so much expenditure and therefore *ipso facto* to provide the system with just that much income. On a second table, with columns and rows similarly labelled, each decision-maker would insert, for each combination of interest-rate and aggregate income, the number of banknotes[1] he would in those circumstances wish to hold. It is a familiar and reasonable assumption that when a man's income is very low he will spend a high proportion, or even the whole of it, on 'food and drink'; when it is very high, he will spend a lower proportion on 'food and drink'. The reverse may be true of his outlay on equipment, but not, perhaps, to such a degree as to reverse the net effect. Taking, let us say, the top row of each of the first set of tables (and for the moment ignoring the other rows), the central office is therefore likely to find that the sum of all the figures entered under some one hypothetical figure of aggregate income is *equal* to that hypothetical figure. Then taking the second row of all the tables, the central office will again find some one column where the figures in the row and column in question add up (approximately) to the number at the head of the column. The ultimate result of such a study of the first set of tables will be a two-column schedule showing, for each hypothesis concerning the interest-rate, some one size of aggregate income. Turning then to the second set of tables, the central office will find, for each combination of a particular interest-rate and aggregate income, a particular *total* number of banknotes which the decision-makers, all taken together, wish in those circumstances to hold. Being informed by the monetary authority what total number of banknotes is to be allowed to be in the hands of the decision-makers, the central office will then inform each decision-maker what the aggregate income and the interest-rate for the coming 'week' will be, and he will then proceed to spend banknotes, on 'food and drink' and on extra equipment, to the amount which he promised that, in the circumstances now laid down, he would do. The banking system will at the same time be instructed by the central office to buy bonds from, or sell bonds to, any decision-maker at the announced interest-rate. On the level of abstraction of this model, and if we think of a small (self-contained) community, it is perhaps not too difficult to suppose that each citizen, knowing the *aggregate* income, could estimate within a little his own share of it. Thus each in drawing up his *ex ante* schedules at the beginning of each 'week' (which would automatically become the effective decision-interval for everybody) would make his conditional promises with *knowledge* of the circumstances in which he would have to fulfil any one of them.

By this model, therefore, we avoid the difficulties raised for theory by the fact that in reality the incomes which individual decision-makers assume or expect in deciding how much to expend on consumption or net investment may add up to more, or less, than the total of those individual expenditures. Such a discrepancy between aggregate expected income and aggregate intended expenditure must mean that some individuals will find that their decisions have been based on wrong premisses. The continuous emergence of such disappointments and the continuous efforts by individuals to adjust their resulting situations must affect the course of events; but in a first approximation we wish to avoid these complications.

---

[1] By banknotes I mean, as explained above, bank balances as well as physical notes.

In our model, aggregate income is composed exclusively and precisely of aggregate expenditure on consumption plus aggregate expenditure on net investment. Aggregate individual saving I define, as usual, as aggregate income less aggregate spending on consumption. Thus it follows that aggregate saving by all individual decision-makers taken together is identically equal to their aggregate net investment, and this is true *ex ante* as well as *ex post*. What, then, will be the effect of a general weakening or strengthening of the desire of individuals to consume out of given incomes? Can it affect the interest-rate? Can the interest-rate in changing have repercussions on the number of banknotes saved per unit of time? In short, is there even partial truth in the proposition that the interest-rate is what equates the aggregate saving-flow and the aggregate net investment-flow?

It is expenditure which, in our model and in other more nearly realistic ones, elicits and measures output. A decision *not to spend* a given number of banknotes is a signal to someone *not to produce* that number of banknotes' worth of output: in our model this signal is given and obeyed explicitly, exactly, and without time-lag. A general weakening of the desire of individuals to consume out of given incomes means that each will enter a smaller figure than hitherto in *each* column of his expenditure array, this smaller figure being composed of a smaller amount for consumption and an unchanged amount for net investment. That total of conditional promises to spend, which matches the heading of the column in which these promises stand, will then be found in a column headed with a *smaller* aggregate income than the one which formerly yielded the solution. Thus, as the outcome of the transformation of individual income-consumption functions to greater frugality, the aggregate income will be smaller, the pace of accumulation (i.e. the identically equal flows of aggregate saving and aggregate net investment) will be unchanged. Nothing will thus have happened to the saving-flow, and any effect there may be on the interest-rate cannot occur directly through that flow. But the smaller total value per unit of time of transactions in general, corresponding to the smaller aggregate income, will require a smaller stock of banknotes to handle it with a given degree of convenience. Out of a given total number of banknotes in existence, more will then have to be held by persons induced by a low interest-rate to prefer them to bonds. In this indirect manner the rate may, to some small extent, be affected.

So far in this paper I have been outlining a general and largely formal theory. Neglecting the plurality of interest-rates, the great variety of 'fixed-interest' securities ranging from treasury bills to Consols and extending to commercial debentures of lower credit standing, and neglecting also the variety of 'legal persons' who lend and borrow, as different in their needs and their legal obligations as private mortals of small wealth and institutional immortals of huge wealth such as banks and insurance companies and charity trust funds, I have been concerned only with the most fundamental questions: why there should be interest, what part of the essential character of economic life it symptomizes; and by what psychic mechanism we can suppose men to choose determinately between such contrasted satisfactions as consumption and accumulation (which turns out, however, to be of minor relevance for us), and between *security* on one hand or *enlargement of hopes* on the other, between cramping but protecting *certainty* and hazardous freedom to fear loss and hope for gain. But now I wish to turn to detail and to factual illustration.

If the yield of a bond[1] is that feature of it which affords a *presumption* that, when all is said and done, the total number of banknotes eventually received by the purchaser of the bond, whether by reselling it or by holding it until redemption. will turn out to be greater than the number he gave for it, then this yield will surely be smaller when the lender's hazard is confined within narrow limits. Now when a security is due to be redeemed (that is, a final payment equal, or about equal, to the amount originally lent is due to be made) quite soon, the holder of it, should he unexpectedly need ready money, will easily find a buyer who feels sure of being able to wait the short time until redemption: the shorter that time, the more numerous such willing buyers will be, the more eagerly they will compete with each other for a small but sure gain, and the nearer, therefore, the market-price of the security will be to the final 'redemption' payment. Thus the yield, we should expect, will be low. By contrast, when the series of payments still due from the debtor stretches over many future years, it becomes impossible to guess what the terms of borrowing and lending will be at such remote dates: so much can obviously happen in the meantime to upset all calculations based upon the present visible situation. Thus as we turn from treasury bills, and bonds already called for redemption within a few months or a year or so, to bonds whose remaining life is still several years, we should expect the yield to rise rapidly at first: a bond with a remaining life of 4 years will have a much higher yield than one of only 6 months. But the power of 'present' knowledge to penetrate the future does not even *seem* to us to be much (and surely is far less than our indolent and unimaginative natures incline to suppose it). Enough hazards can crowd themselves into a remaining bond-life of 5 or 10 years to make us care little about the additional ones belonging to a life of 20 or 30 years, and we should accordingly expect the rate at which yield increases with increase of bond-life to slacken rapidly beyond the earliest years.

---

[1] It will be remembered that we are using 'bond' to mean any promise-to-pay by an undoubtedly sound debtor.

But if differences in length of life account for *relative* bond yields, what accounts for their *general level* at any instant? The ultimate refuge from market uncertainties (unless faith in the value of money is lost in a galloping inflation) is money. Let us imagine two situations such that each individual attaches the same degrees of doubt or assurance to given hypotheses in the one situation as he does in the other (*different* individuals perhaps differing from each other very widely in both the one situation and in the other); two situations, that is to say, which are identical in regard to the state of mind, knowledge, set of hypotheses entertained with given degrees of uncertainty, and attitude to such uncertainty, of every individual; and let us suppose them also identical in regard to the quantity existing of each form of wealth, and its distribution; except that the quantity of *money* is much larger in situation *B* than in situation *A*. Then it is plausible that the relative scarcity, in situation *A*, of that form of wealth which gives shelter from whatever market uncertainties exist in the minds of the various individuals will make *banknotes dear in terms of bonds* in that situation; that is, the *yield* of bonds will be higher in situation *A* than in situation *B*. If the monetary authority could in reality effect a transition from *A* to *B*, could, that is to say, largely increase the number of banknotes in people's possession[1] without thereby changing their expectations, it is plain that the price of bonds in terms of banknotes could be raised, in other words the yield of bonds could be lowered. It is plain also that though in reality any act of largely increasing the quantity of money would inevitably change people's expectations, some changes of expectations are easily conceivable that would strongly assist the reduction of interest-rates. Indeed, in a 'cheap-money drive' nothing succeeds like success; for if once a consensus of market opinion can be swung, by words and deeds of the monetary authority or its masters, to the view that bond prices will go yet a little higher, the resulting preponderance of orders to buy will push them that little higher, and justify the suasion that induced the buying; and thus step by step the monetary authorities can drive bond prices up and up: *possunt quia posse videntur*.

How, then, are we to explain the ultimate failure of the attempt, begun in the autumn of 1945 and backed by all the resources of modern government used with complete determination, to drive down the long-term interest-rate from its starting-level of 3 per cent. per annum to 2½ per cent. or even lower, *and to establish it there as a respected permanency?* To understand this failure we must consider the nature of the temporary success that was achieved.

---

[1] In the bank account or pocket.

The rate of interest is, of all prices, the one most inseparably bound up by the logic of its very nature with expectation and uncertainty. These are, indeed, the main reason for its existence. This is true at all times, but nowadays for two reasons this 'dynamic' character of interest-rates has come to be able to dominate all other aspects completely: first, the disturbed and 'experimental' character of the times makes the possibility of dramatic changes more real and insistently present to people's minds; and, second, the scale and incidence of taxes in Britain to-day makes *capital gains* immensely more attractive than what is conventionally thought of as 'income' from securities. Capital gains are untaxed, so that while 'interest' payments of £1 a year can to a rich man be worth as little as 6*d.* a year tax free, a capital gain of £1 is worth £1. This factor of taxation immensely increases the sensitiveness of long-dated gilt-edged security values to uncertainty and suggestions of change in trend or policy. For a rise in the yield of irredeemable stocks from, say, 2½ per cent. per annum to 3½ per cent. per annum, such as occurred in Consols between November 1946 and April 1948, implies a capital loss sufficient, with tax at 9*s.* in the £, to wipe out some 16 or 17 years' net income from 'interest' payments, and correspondingly more as surtax increases. This state of affairs must mean that for most owners or free administrators of wealth, other than 'immortal' institutions not subject to tax, the hope of capital gains and the fear of capital losses completely dominates their outlook to the exclusion of the question of yield.

In short, bond prices were able to go on getting higher and higher, bond yields correspondingly lower and lower, so long as enough people were able to believe that they would go on a little longer. *But they could not stand still at an extreme level and stay there.* Once the rise had been checked, the spell broken, fears of capital loss took the place of hopes of gain. A bond whose price is *high and rising* can look profitable: one whose price is high and *not rising* is a mere peril. The rout, once started by the fuel crisis, carried Treasury 2½ per cent. stock down from 100 to 83 in 9 months.

These ideas and facts are illustrated in the accompanying charts. In Chart 2 the quantity of money, defined as the total deposits of the London Clearing Banks plus estimated notes and coin in circulation with the public,[1] is shown as a heavy continuous line *M*; the 'notes and coin' component of the quantity of money is shown as a light continuous line *n*; and the yield of

---

[1] From *Monthly Digest of Statistics* (H.M.S.O.).

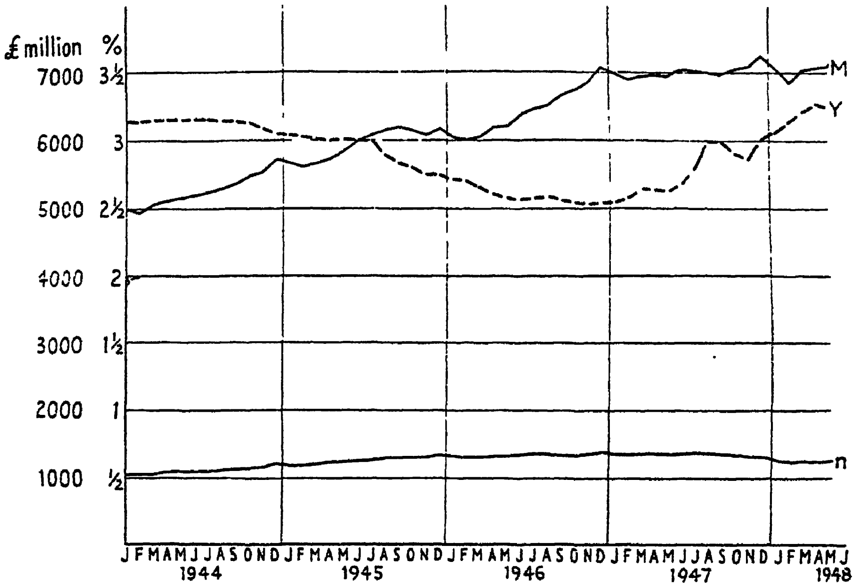

CHART 2

M is total deposits of London Clearing Banks plus notes and coin with the public. *Y* is yield of 2½ per cent. Consols. *n* is notes and coin with the public.

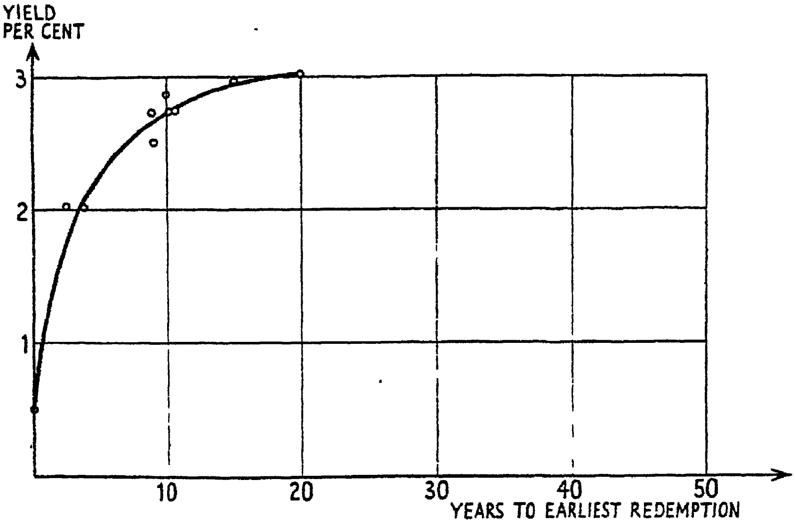

CHART 3. Yield as a function of remainder of life at 7 September 1945

Quantity of Money: £6,200 *Mn.*

2½ per cent. Consols is shown as a line of heavy dashes *Y*. In mid-1945 the quantity of money had been increasing (with only seasonal and other minor interruptions) ever since the beginning of the war; the yield of Consols had been rather steady for some years at a little over 3 per cent. The 'cheap-money drive' began in September 1945, and between then and November 1946 the quantity of money was increased by some £650 million or 10 per cent. Moreover, the whole of this increase was in deposits and none of it in notes and coin. If the latter (as seems likely) are a fairly constant proportion of the money required for convenience in handling transactions, all the extra money must have been held as an alternative to other ways of storing wealth. (Between February and December 1946 the quantity of money actually increased by £1,000 million, but part of that increase is seasonal.) The yield of 2½ per cent. Consols fell, between those dates, from 3 per cent. per annum to 2½ per cent. In the spring of 1947 the fuel crisis shook the belief in the Exchequer's power and willingness to go on supporting and shouldering up the bond market: failure of belief in a *continuing rise* meant, inevitably, a rush to get out of gilt-edged at the perilous heights they had reached. Although the quantity of money continued, though very gently, to *increase*, the yield of Consols sprang back within 6 or 7 months to 3 per cent., and in another 8 months to 3¼ per cent., higher than it had been when there was £1,000 million less money. But is this perhaps true only of irredeemable stocks? Let us look at Charts 3, 4, and 5, plotted from data given in the accompanying table.[1]

Yield as a Function of Number of Years to Earliest Redemption, at Three Dates

			7 Sept. 1945		*14 Nov. 1946*		*13 Aug. 1947*	
			Years	*Yield*	*Years*	*Yield*	*Years*	*Yield*
Conversion	3%	48/53	2·5	2·020	1·3	1·380	0·5	1·130
Nat. War	2½%	49/51	3·9	2·016	2·7	1·390	2·0	2·070
Defence	3%	54/58	8·9	2·720	7·7	1·940	6·9	2·875
Nat. War	2½%	54/56	9·0	2·500	7·7	1·870	7·0	2·725
War Loan	3%	55/59	10·1	2·738	8·9	1·988	8·2	2·912
Savings	3%	55/65	10·0	2·880	8·8	2·042	8·0	2·980
Funding	2½%	56/61	10·6	2·742	9·4	1·921	8·7	2·920
Savings	3%	60/70	15·0	2·970	13·8	2·234	13·0	3·046
Savings	2½%	64/67	..	..	17·5	2·184	16·75	2·925
Savings	3%	65/75	20·0	3·012	18·8	2·300	18·0	3·046
Treasury	2½%	75 or after	..	..	..	2·500	..	3·075
Consols	2½%		..	2·830	..	2·525	..	3·029

In each of these charts the curve relates the *yield* of bonds to their *remainder of life*. By the remainder of life of a bond I mean the time which

---

[1] I drew all these curves freehand. Any more elaborate procedure would, for my present purpose, have been a waste of effort.

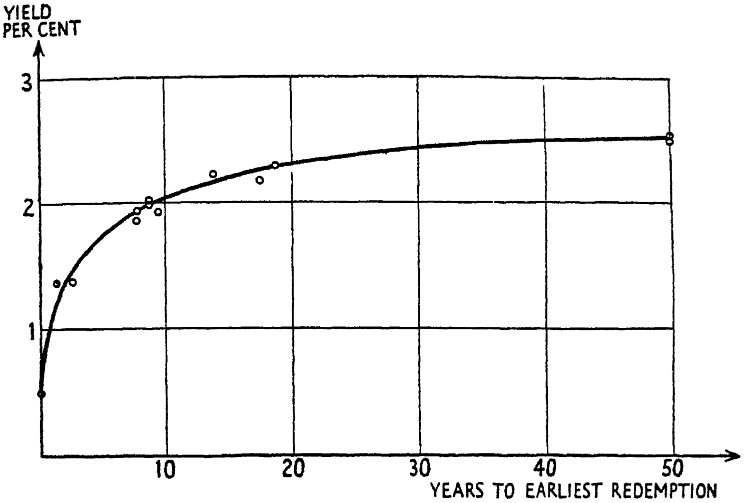

Chart 4. Yield as a function of remainder of life at 14 November 1946

Quantity of Money: £6,850 *Mn.*

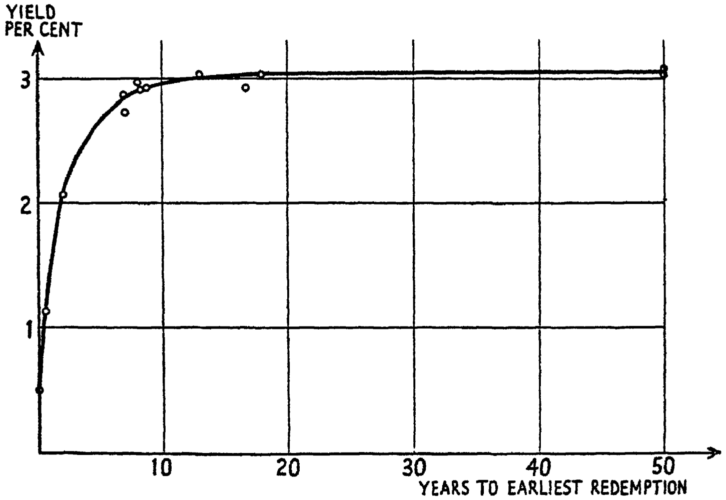

Chart 5. Yield as a function of remainder of life at 13 August 1947

Quantity of Money: £7,000 *Mn.*

has still to elapse before the debtor's final payment. For irredeemable securities, viz. 2½ per cent. Consols and Treasury 2½ per cent. stock, this life has been conventionally taken as 50 years. To represent a zero life I have used treasury bills. The other securities for which points are plotted on these charts have each a *range* of dates within which 'repayment' will be made, instead of a single date. For all these the *earliest* date of the range has been taken.

As we should expect from the foregoing analysis of the nature of interest, the slope of the curve in each chart is steep near the origin but becomes rapidly less steep[1] beyond 5 or 10 years. However, our main concern is a comparison of the shapes and positions of the curve at three different dates. Chart 3 shows yields at 7 September 1945, which I take as the start of the 'cheap-money drive'; Chart 4 shows how, by 14 November 1946, the whole curve had been pushed bodily downwards, and Chart 5 shows how by 13 August 1947, after the convertibility crisis, the curve had gone up again beyond its level at September 1945. In Chart 6 all three curves are plotted together for easy comparison.

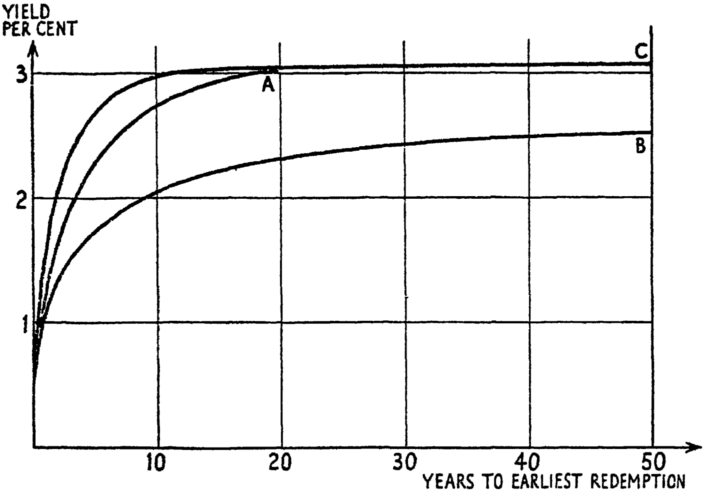

CHART 6. Yield as a function of remainder of life

	Curve A	*Curve B*	*Curve C*
Date	7 Sept. 1945	14 Nov. 1946	13 Aug. 1947
Quantity of Money .	£6,200 *Mn.*	£6,850 *Mn.*	£7,000 *Mn.*

These violent movements ought to shock anyone who still believes that the rate of interest depends on the 'credit standing' of the government, or on the degree of thriftiness or extravagance of the people. They merely magnify and dramatize the fact that interest is a manifestation of uncertainty, and in highly uncertain times it will betray its nature by the scale and momentum of its changes.

LONDON.

---

[1] See pp. 114, 115 above.

TYPES OF COMPETITION AND THE THEORY OF EMPLOYMENT

By B. R. WILLIAMS

In recent years the main changes in the presentation of the elements of economics have been connected with the theories of imperfect competition and 'effective demand'. It has often been suggested that these theoretical developments, and the alacrity with which they were accepted, were due to an awareness that the older theories were not adequate to explain the facts of economic life, and that in being modified to 'explain' them internal inconsistency resulted. It has been suggested, too, that these new theories strengthened the theory of value and broke down the old division between value and monetary theory—that on the one hand the theory of imperfect competition strengthened the theory of the firm and so the theory of value based on the preferences and reactions of individuals, while on the other hand Keynesian theory, in breaking away from the 'traditional lines of regarding the influence of money as something separate from the general theory of supply and demand' by explicitly introducing the marginal principle into the money market and a 'fundamental psychological law' of consumption, linked the theory of employment with 'the fundamental theory of value'.[1]

So far, however, there has been no synthesis between 'value' theory and the theory of employment. Lord Keynes was concerned with working out a theory of the demand and supply of output as a whole. Though he asserted that his theory was linked up with the 'fundamental theory of value', it ran mainly in terms of aggregative functions (e.g. entrepreneurs collectively maximize an aggregate profit) very loosely related, if at all, to the individual functions of value theory. The 'General Theory' grew out of a theory of money in which the changes in the price-level determined profits and losses, and therefore output, without reference to individual households and firms,[2] and no real attempt was made to explain the form of the individual functions that would correspond to the aggregate functions. As a result, it will be argued in section 4, Keynes made use of certain conclusions of value theory derived from static individual functions the form of which is inconsistent with his own assumptions. Despite this, or perhaps because of it, Lord Keynes attempted to prove that there need not be full employment even in a perfect competition equilibrium—an equilibrium which is not explicable in terms of his aggregative functions.

---

[1] J. M. Keynes, *The General Theory*, pp. vi, vii.

[2] Cf. J. A. Schumpeter, *Business Cycles*, p. 43.

The avowed intention of imperfect competition theory was to explain the ground in between perfect competition and pure monopoly. But although that was not fully possible so long as the theory of value was treated as the 'analysis of the output of a single commodity considered separately',[1] the attempt was made in terms of the shape and position of the average revenue curves of individual firms. Any attempt to analyse imperfect competition from the basis of a single firm with a given average revenue curve somewhere between the limits of infinite and unit elasticity has to be so purely formal that the fundamental features of markets which are not perfectly competitive or perfectly monopolistic are left unexplained. Nevertheless, in suggesting that for analytical purposes it is possible to move easily between the limiting cases of perfect competition and perfect monopoly, these new theories of imperfect competition, whether they stem from Mrs. Robinson's tacit assumption that every firm but one is in equilibrium or from Professor Chamberlin's assumption of symmetry,[2] have placed a premium both on the use of perfect competition as the standard case and on the development of a value theory 'pure' enough to facilitate the maximum use of plane geometry.

Yet an examination of the abundant supply of text-books will show that insufficient attention is paid to the assumptions of this perfect competition. To take one important example, it is rare to find a distinction between the assumptions, first, that the current market-price is the planning-price of firms and households, and, second, that the individual cannot by his own action affect the price. For the usual diagrammatic treatment of perfect competition both assumptions are necessary, but for Keynesian theory the first is not even a possible assumption. We can find a similar lack of care in the statement of the conditions of free entry. The fact that 'free entry' is dependent on a positive reaction of saving to the rate of interest is in itself an indication of the difference between the 'perfect competition' of 'value' theory and that of 'employment' theory. That the implications of free entry have been neglected may be explained by the facts that, first, in considering separately the output of a single commodity, 'marginal doses' may be drawn in from 'elsewhere', and, second, that in a purely formal treatment of competition, 'free entry' and 'imputation' are alternative ways of making 'profit' disappear. That *apparently* the same perfect competition is capable of producing both an optimum allocation of resources when approached through individual functions, and unemployment equilibrium when approached through aggregate functions, indicates the need for much greater care in the statement of the problem. Just why this change of *method* should produce such a different result is rarely made clear, and the student is not to be blamed if he concludes that value theory is no more than a traditional pedagogic instrument.[1] More serious is an apparent inconsistency between the operation of pricing mechanism in the cases of 'value' and 'employment', despite the alleged abolition of the inconsistent treatment of money as between value on the one hand, and international trade and employment on the other.

---

[1] Joan Robinson, *The Economics of Imperfect Competition*, p. 6.

[2] v. Triffin, *Monopolistic Competition and General Equilibrium Theory*, chapters 1 and 2.

The difference made to the co-ordinating effects of the pricing mechanism by different assumptions about expectations is so great that it seems useful to distinguish between two main types of competition. First there is the type of competition implicit in the theory of optimum allocation. According to this theory, perfect competition in bringing about a no-profit equilibrium gives, after allowance for 'smoke nuisance' and so on, the maximum satisfaction to consumers at the least cost. For such a theory, as usually expounded with supply and demand curves conceived as independent of the price to which they lead, the price-expectations must be static—i.e. plans must be made as if the present market-price will continue. However, as will be seen in sections 3 and 4, the need for static expectations is independent of the form of exposition. Furthermore, it is necessary for optimum-allocation theory to show that full employment is produced by the operation of the pricing mechanism, and that demonstration involves the assumption of static expectations. Accordingly we call this first perfect competition 'static perfect competition'.[2]

Second, there is a theory of how competition acts as a spur to efficiency ('compels progress'), as a way of eliminating the unfit, and sometimes, as, for example, in Marshall, as a means of getting capital into the hands of those most capable of using it. Such a theory of competition, which we term 'dynamic competition', leads to the conception of the entrepreneur as a leader and transformer of economic life, in contrast with 'static competition' which leads to the conception of the 'consumer as king'. With 'dynamic competition' it is not possible to treat entrepreneurs[3] as simply adapting themselves to current market-prices. It is necessary to take account of divergences between expected and current prices, which means not only that there will be variations in marginal cost because of the differing expectations of different individuals, but also that, by virtue of non-static expectations and the part played by credit creation in the finance of innovation, dis-equilibrating monetary movements become possible. It will be argued that dynamic competition cannot be 'perfect', and that even a definition which runs solely in terms of the inability of a buyer or seller to influence the price by his own single-handed action (and which by itself carries no proof of an implicit tendency to equilibrium) is not always applicable. As, in addition, the sort of expectations assumed in Keynesian theory are explicable only in terms of conditions which are not at all points amenable to marginal analysis, it will be argued that there is no meaning in the 'Keynesian' proof that there can, and generally will, be unemployment even in perfect competition equilibrium.

---

[1] Cf. P. Sraffa, *Economic Journal*, Dec. 1926, p. 535.

[2] Cf. N. Kaldor, 'A Note on the Determinateness of Equilibrium', *Review of Economic Studies*, vol. i.

[3] Throughout this paper the term 'entrepreneur' is confined to conditions of dynamic competition. In static competition firms and households simply adapt their plans to current market-prices, but in dynamic competition entrepreneurship is involved primarily in the introduction of new production functions, and secondarily, in a situation disturbed by change, by profit 'maximizing' activities which cannot be based on a relation between present prices and costs. Thus static analysis, as Prof. Schumpeter has often stressed, cannot be extended by the introduction of the 'cost of entrepreneurship or risk-taking' to deal with the dynamic problem. Cf. also below, p. 126, n. 2, and sections 3 and 5.

1

Perfect competition is commonly defined as that state of affairs where there is no artificial restriction upon movement of factors of production from occupation to occupation in search of the highest reward, and where no single unit of control (firm or household) can by its own action appreciably affect the price of the thing bought or sold. For the purpose of this and the following section, where we shall deal with 'static perfect competition', we make the further assumptions that firms and households take the ruling market-price as the basis of their plans, that production functions are given, and that money is but a medium of exchange.[1]

Does it follow from these assumptions that any divergence between price and average cost will induce firms to enter or leave a particular industry till, by that entry or exit, the divergence is abolished? As we are concerned with examining the conditions of free entry it is useful to take as the starting-point the procedure in Meade's *Economic Analysis and Policy*, where it is argued that such a movement will take place, even where there are fixed capital goods—that although the transference of fixed capital goods must take time, nevertheless it will in the end be achieved as the machines in the loss industries are not renewed and the depreciation funds are invested more profitably.[2] As this movement is not made to depend on the provision of new capital, but simply on the search of existing factors for the highest reward, i.e. on the reinvestment of old capital and its embodiment in a different material form, it seems to depend on a conception of capital as 'an entity capable of maintaining its quantity while altering its form, and by its nature always drawn to those forms on which, so to speak, the sun of profit is at the time shining'.[1] The general criticism of this conception has been made by Wicksell,[2] but as the coexistence of optimum-allocation and unemployment-equilibrium theory must be explained in some measure by different assumptions about the market for loans, it seems useful to go over the matter again, and as an aid in bringing out the implications of reactions in the loan market we suppose that there is no market for loans.

---

[1] The relation of perfect competition to unemployment will be examined in section 3. It is not sufficient to rule out unemployment as inconsistent with a postulated equilibrium, as is, for example, the procedure of A. L. Meyers in *Elements of Modern Economics*. In *Economic Analysis and Policy* J. E. Meade supposes that any deficiency of effective demand is remedied by State action, but he gives no hint as to how account is to be taken of such action in the determination of relative prices. The confession of the need for this State action also mocks at his doctrine that under perfect competition the profit interest of producers leads to maximum production.

[2] pp. 103–4.

Let us postulate a 'long-period' equilibrium in which for each firm price equals minimum average cost. Included in that cost will be a return to the owner of the firm (the capitalist) which will just induce him to maintain the capital of that firm. At this point let us take as given the initial capital of each firm, and suppose for the first part of this section that this capital can be added to or depleted, not by the decision of the capitalist as householder, but only by trading—i.e. only by profits and losses incidental to an autonomous change in the direction of consumers' demand. If a firm makes profits because of such a change, the profits are retained by the firm so long as they can be turned into additional profits, i.e. so long as, in general, price exceeds long-period marginal cost. If a firm makes losses, not only can no funds be obtained from outside, but depreciation funds are not maintained at the expense of the capitalist as householder. Where the return to the householder is reduced, that will be because current gross receipts are not sufficient to make the original return to the householder possible. This appears to be the clearest way of separating the accounts of households and firms, given the assumption of no borrowing.

We have postulated an autonomous change in consumers' demand which disrupts a previously existing position of long-period equilibrium. As a result, some firms will receive more from the sale of their produce than is necessary to maintain their former scale of operations, while others will receive less. Capital values will rise in the 'profit industries' and fall in the 'loss industries'. But this means that firms in the loss industries, some of which are, in accordance with the mechanism mentioned above, supposed to migrate to the regions where the sun of profit is shining, will have a smaller command over productive resources. The longer firms make losses, i.e. the more specific their equipment, the farther away the time when that equipment would have been renewed, and the greater the proportion of fixed to total cost, the greater will be the loss of capital.

---

[1] A. C. Pigou, *Economic Journal*, 1935, p. 239.

[2] *Lectures on Political Economy*, vol. i, pp. 200–6. See also Hayek, 'The Maintenance of Capital', *Economica*, 1935.

Suppose, first, that there are constant returns to scale of plant. Then when profits are earned it will pay existing firms to add to plant as long-period marginal cost is less than price. And so long as long-period marginal cost is less than price profits will be accruing to enable that expansion. The profits of the fortunate firms, which have benefited from the autonomous change in consumers' demand, will be a source from which the finance necessary to purchase the new equipment can be derived. Thus these abnormal profits provide at once an impetus to expand and the means of doing so. Indeed, the necessary new capital (offsetting which is a destruction of capital in the loss industries) could be provided entirely out of profits, and in the absence of the transfer of firms from other industries would be so provided.

That there must be a destruction of capital in the loss industries is clear. Professor Knight has pointed out that where a breath could transform a brick factory building into a railway yard or an ocean greyhound, there would be no need for economic activity or economic science,[1] and in any case the assumption that the ruling market-price forms the basis of plans rules out the use of re-contract or provisional *tâtonnements* to make, even when there are no fixed factors, the movement of factors and the formation of equilibrium prices precede exchange. The fact that there are fixed factors means that capital losses inevitably follow a change in demand. Furthermore, these fixed factors, or frictions, are a pre-condition of the operation of the pricing mechanism. For if there were perfect mobility the emergence of a difference between price and cost in one sector would cause such a flood of resources there as to ensure losses for all.[2] Nor could the flow back to the now prosperous deserted sectors be orderly, for with static expectations and perfect mobility (which implies that no producer has a preference or special competence for one industry rather than another) there would be nothing within the pricing mechanism as such to make it possible for producers to choose a profitable transfer.[3]

[1] *Risk, Uncertainty, and Profit*, p. 34.

[2] Judged by result, actions based on the current price without reference to the reactions of other producers, where they are significant, is evidence of lack of foresight, though in stationary conditions only such action is consistent with perfect foresight. The difficulty lies in treating expectations as independent of the course of events, but this difficulty cannot be great as soon as it is realized that the assumptions of static expectations and rational action can only be legitimately combined in quasi-stationary conditions.

[3] This inability to choose a profitable transfer applies both to the immediate and farther-off future unless there is some force at work to reduce the fluctuations over time. But with static expectations and perfect mobility only the progressive destruction of capital due to losses would do that. For in the flow back to the 'now prosperous deserted sectors' we cannot rely on the law of large numbers to bring it about that too many firms will not go to one industry or another, for it is likely that industries will have need for widely varying numbers of firms.

The avoidance of such a breakdown of the pricing mechanism is due to lack of mobility, and to firms having, at any one point of time, differing degrees of mobility. The unequal incidence of loss, over time, is a necessary condition of a movement to equilibrium, and this implies that the relative size of firms, and the ownership of capital, will be altered by a change in demand. But this will not simply be a redistribution of the old capital—part of that will have been destroyed, and part, at least, of the new capital will have been created out of profits.

Just as in the industries which profit from the change in demand it is possible to get equilibrium from an increase in the size of existing firms, so in the industries making losses it is possible for losses to be removed entirely through the contraction of existing firms. However, some transfer of capital to the profit industries will take place, and, as it is possible to set up any scale of plant without loss of efficiency,[1] those firms which transfer while profits exist will earn surpluses and expand. This transfer will increase the tempo of adjustment in both types of industry, and in reducing the amount of new capital to be created out of profits will reduce the cost to consumers of the change in the direction of demand. However, though the transfer of capital from loss industries will benefit consumers, that transfer, given both constant returns to scale and static expectations, is not necessary to obtain a no-profit equilibrium. The indeterminacy of the division of total output between firms is then immaterial.[2]

An objection to the assumption of constant returns to scale is that there are often indivisible services which lead to alternating stages of increasing and diminishing returns to scale of plant. For this reason the assumption most open to objection as lacking in generality is that firms can set up, on however small a scale, without loss of efficiency. Ease of entry and the possibility of smooth plant expansion will be the less the greater are the indivisibilities. Let us suppose that there is some factor which makes the long-period average cost curve (the interplant curve) for each firm U-shaped. Taking first the firms already in the industries which have profited from the change in demand, it can be seen that they will gradually expand plant as surpluses accrue, though not continuously if there are indivisibilities, till long-period and short-period marginal costs are equal to price. But since long-period marginal cost is rising, the expansion will cease at an output where price is greater than average cost. A no-profit equilibrium cannot be reached through an expansion in size of existing firms.

---

[1] This does not mean that the stabilizing effect of different degrees of mobility will be overcome.

[2] The definition of perfect competition on p. 124 above included the assumption that no firm can by its own action appreciably affect the price of the thing sold. But if the number of firms did not change when demand increased substantially one or more firms might be capable of affecting the price. Thus, if the number of firms in the profit industry is not increased by a transfer of firms from loss industries, there could be a clash between the above-mentioned assumption and the assumption that the ruling market price is the basis of plans. This could however be a short-period inconsistency only, for if the size of firms is such that there is a clash between the assumptions of static expectations and rational action, and the firms act rationally, the failure of firms to use profits to destroy profits must, because of constant costs, lead to an entry of new firms.

Firms in industries making losses, or in industries in which there is no gap between prices and costs, will thus have an incentive to move into the industries which have benefited from the change in consumers' demands, even after existing firms in these industries have maximized long-period profits. Such an influx will be necessary to bring a no-profit position in the expanded industries. But as a transfer of firms from industries undisturbed by the change in utility functions will create there a gap between prices and costs which cannot be filled by the remaining firms, we can consider the achievement of a no-profit equilibrium throughout industry in terms of a transfer of firms from industries which are making losses. Now given the U-shape of the long-period average cost curve for each firm, the greater the rise in price above the minimum point of this curve the smaller the capital with which a move is possible. But the greater this rise in price the greater the loss to those not in a position to transfer quickly. Furthermore, it does not follow that a firm which does transfer while price is greater than minimum long-period average cost will be able to reap profits. If, given the scale at which it is able to enter, price is equal to its minimum short-period average cost, then even though it could expand through the range of descending long-period average cost it will not have the resources for such a profitable expansion of plant, while any expansion in the output of the industry will bring price below the average cost from operating with present plant. For this reason there may be a situation where original firms, as well as those able to come in with sufficient capital to reap profits, have resources with which to expand but no incentive to expand since expansion has proceeded till long- and short-period marginal costs are equal, whereas other firms possessing every incentive to expand, long-period marginal cost being less than short-period cost, are unable to expand because price is not greater than short-period average cost. In addition, even if the optimum size of firm was smaller in the profit industries than in the loss industries, it is conceivable that the reduction in the number of firms in the loss industries, by death or by transfer, might restore profitability there before profits have been destroyed elsewhere. Thus, given the assumptions stated on p. 125, it is a necessary but not a sufficient condition for the achievement of a no-profit equilibrium throughout that the destruction of capital caused by the change in the direction of consumers' demand be offset, after allowance is made for some increase in capital from profits, by a diminished need for capital in the new conditions of demand. How much capital-saving the new utility functions need to be cannot be indicated except in terms of the varying age of equipment in all firms concerned. In contrast, then, with the case of constant returns to scale of plant, there is no mechanism to ensure the use of profits to destroy profit, even in perfect competition. It follows that, except where there are constant returns to scale of plant, the limitation of the power of firms to purchase factors to their receipts from sales means that there is no mechanism to ensure a no-profit equilibrium.

So far we have taken as given the initial existence and size of firms, and have assumed a relation between households and firms that enables a separation of their accounts appropriate to the assumption of no borrowing. The problem of why, and to what extent, proceeds which could be used by capitalists as householders are used in firms has been ignored. It is customary to assume that each household has a scale of preferences. In this way a household with a multiplicity of objectives can be treated as maximizing a single magnitude—utility—and thus can be treated in the same way as firms which are assumed to pursue a single objective—profit. However, the assumption of rational action must be used with care since in the case of the owner of the firm ('the capitalist') the maximization of profit and the maximization of utility are not necessarily consistent objectives. Where consumption and production are treated separately—the first in terms of households, the second in terms of firms—this difficulty does not always intrude, but the impossibility of any hard-and-fast separation is indicated not only by the relation between wants and activities, as stressed by Marshall, but also by the fact that supplies of factors cannot be determined apart from the preference system of households. Thus the number of hours a labourer will work at particular wage-rates is determined by the choice between more income and more leisure, while we cannot know how much saving is involved in the equalization of the present and future marginal utilities of the goods saved without knowledge of households' preference systems. In the case of the owner of the firm, if the maximization of profit is always to be consistent with the maximization of utility, his preference system must be such that any choice between more income and more leisure is independent of his income.

A household in maximizing utility does so within a practicable field of choice. If a labourer has the choice between working for eight hours a day at a particular wage and not working, the problem of a balance between more income and more leisure each day does not arise. In a situation where competition maintains equality between prices and costs. only the maximization of profit is consistent with the survival of the firm.[1] In this sense the capitalist's field of choice does not include a choice between more income and more leisure—competition compels profit maximization. In the case considered above, where there was no mechanism to ensure equality between prices and costs, the capitalists' field of choice becomes greater. There is then no guarantee that profits will be as large as possible in all cases. A further difficulty exists in that there is no reason to expect uniformity of behaviour in the face of either surpluses or deficits in firms. For example, firms in the loss industries can no longer be treated as automatically limited by the extent of losses. To know the reactions of different capitalists to even the same loss—to know whether they will maintain their consumption at the expense of the depreciation allowance, reduce consumption to the level at which it could be permanently maintained during the life of the existing investment, or maintain their depreciation allowances, i.e. save, against the transfer to a more profitable industry—we must look to their respective indifference systems. Because of the absence of borrowing at a common rate there will not even be uniformity in the 'cost' of capital, and hence no uniformity in costs of production. Similar difficulties exist for firms making profits, so that, even apart from costs of movement, there will be no common rate of profit, but only a movement to a maximum in the case of each individual.

The difficulties outlined in the previous paragraph apply also in the case of constant returns to scale of plant, so that in this case too the 'conclusions' of perfect competition cannot be reached without consideration of the market for loans. Even in the case of a simple shift in demand, it is not possible to include the quantity of capital as a known, still less to assume free entry without first showing that reactions in the market for loans will make it possible.

2

The introduction of borrowing and lending brings two strong influences to bear on the rate of profit. First, the existence of a market rate for borrowing and lending sets a standard to which there is a profit interest in conforming. Previously, the very condition which makes borrowing and lending profitable—that there are people whose valuation of wealth in hand and in prospect are so proportioned as to give the present a relatively higher place in their scales of preference than it occupies in that of others—prevented there being any one rate at which the expansion of firms would stop. But with a market for loans in existence, the willingness of any capitalist to continue saving at the current rate of interest does not imply that the rate of return on his investment would be pushed lower than that of another capitalist who would not save at that rate.

---

[1] Unless the psychology, and the means, of the capitalist are such that maximum satisfaction implies paying for playing at business. This situation should not be confused with cases which can occur in dynamic competition where the relation between the prevailing rate of interest and the past and prospective returns is obscure; cf. Keynes, *The General Theory*, p. 150.

Second, firms which can borrow will not be in a position of having an incentive to expand but no resources with which to expand. But in the situation we are discussing firms will only be able to borrow if saving is an increasing function of the rate of interest. In a position of long-period equilibrium the rate of interest must be such as to make saving zero. With the change in consumer demand there will be a demand for new capital which can be satisfied only if a greater rate of interest makes saving positive. If it does, firms which are capable of expanding through a range of decreasing cost would not only be able to make profits, once past the point where price is equal to average cost, but would be able to make a greater profit per unit of output than established firms operating through a range of increasing cost. If, on the other hand, saving does not increase, there is nothing other than the possible investment of capital gains on the part of firms in the profit industries to bring an adjustment, and, if we rule out formal imputation which substitutes a factor profit for 'profit', this cannot ensure a no-profit equilibrium. Without a positive reaction of saving to the rate of interest 'free entry' is not possible, and the assumptions set down at the beginning of section 1 are not sufficient to ensure a no-profit equilibrium.

We suggested above that one reason for the neglect of the conditions of free entry was to be found in the misuse of partial equilibrium analysis. We have taken one case which is misleading since the 'given' factor is not independent of the path of adjustment. Of particular danger in the case of free entry is that a solution worked out for one market in isolation may be implicitly extended to cover all markets, as if solutions reached in this way for each single market in turn could be added to reach a legitimate total. Writers who consider the maximization of profit, on the basis of the perfect elasticity of the supplies of factors to the individual firm, before considering supplies of factors as a whole, are in danger of committing this second error. The implication of such a procedure will be considered in terms of the reactions necessary for the achievement of a no-profit equilibrium.

In examining the adaptation of firms to profit opportunities there are two aspects to be considered: the plans of individual firms on the basis of given data, and the effects of attempts to carry out such statically established plans on the data. On the importance attached to these two aspects of the adjustment will depend the sort of analysis which will be used. This is illustrated by the two different types of partial equilibrium analysis. With the first, the equilibrium conditions are defined on the basis of given costs, and the adaptation assumed to take place. With the second, allowance is made for the effect of the adaptation on the 'given' conditions, and although it is impossible in this way to explain the change in data for the partial adaptation, the method is a very useful way of translating the result reached by general equilibrium analysis. The implication of the first procedure is that the conditions of equilibrium are independent of the actual adjustment. For example, if there is only one size of firm consistent with a no-profit equilibrium, the 'proof' that this position will be brought about is that, on the basis of a further assumption about the elasticity of factor supplies, the profit-maximizing activities of producers make it impossible to get equilibrium at any other position. But even where the ruling market price forms the basis of plans, since this means that during a process of adjustment each producer is working with an imaginary demand curve independent of the actual demand curve of consumers, the path of adjustment must not be ignored. For if the supply of factors to any firm, or to any potential firm, is perfectly elastic at a given price, and the price of the product is such as to give profits from its manufacture, what is to prevent a veritable flood of resources to the places where profits exist? That is, what is to prevent the market demand curve from intersecting the cost curve of the industry —the sum of the cost curves of the individual firms—well to the left of its minimum point? And what then is to prevent a further disequilibrating movement in the other direction? With any trial-and-error movements in which some error is involved the plant existing at any time is partly determined by what plant was in the preceding period. Where plant is durable and price-changes are not anticipated, the short-run cost curve need not be the lowest for that level of output, and will not be one of the family of curves on which the long-run curve is based. In addition 'excessive' movements must affect the supply of capital needed for adaptation, for even if existing prices were still taken as the basis of plans, the power to provide capital could not be unaffected by such huge losses.

In fact, the equilibrating movement postulated in the first procedure is that plans are made not on the basis of the current price but of the 'equilibrium' price.[1] If cost curves were given independently of disequilibrium, then the equilibrating movement must take place through the effect of the production from new firms on the price of the product. But a 'free entry' derived in this way is incompatible with the achievement of an equilibrium. For the implication is that the assumption made for the firm, that the supplies of factors to it at the prevailing price are perfectly elastic, can be extended to the new firms as a group, and even to the industry, and it is precisely this extension which prevents there being any stability conditions—that in each market there exists a price to equate supply and demand in such wise that a rise in price will make supply greater than demand, and vice versa.[1] When we are considering the case where, through a change in demand, an increase in the output of one industry is taking place at the expense of another, account must be taken of the coefficients of production in the affected industries and the consequent effect on the prices of factors. But, as we have seen, even if the coefficients are the same, the movement to an equilibrium which excludes profits is dependent on the formation of new capital. Where it is assumed that the supply of factors to the industry is perfectly elastic at the current price, the case of capital cannot thus be reduced entirely to the drawing in of marginal doses from other places—some at least must be newly called into being. And if the supply of new savings is perfectly, and instantaneously, elastic at the current rate, the limit to an expansion through a fall in the price of the product must come too late.

---

[1] For the logical difficulties implied in the concept of perfect foresight see Knight, *Risk, Uncertainty, and Profit*, chapters 5 and 6.

The second procedure is not inconsistent with the existence of stability conditions, as account is taken of the effect of the adaptation on the given conditions. Thus the effects of any differences in the coefficients of production as between the expanding and contracting spheres of industry on the relative prices of factors, and of the inconsistency, during a period of adjustment, of statically established plans, are included in the analysis through a decision on the relevant 'given' conditions at each stage of the adjustment. In considering free entry one of the indispensable conditions is the availability of money capital for new firms, and that implies the existence of new savings. That is to say, the abnormal rate of return in the profit industries must induce saving if profits are to be destroyed. Even in the case of constant returns to scale of plant, the movement to a no-profit equilibrium implies the creation of new savings, as, given the assumption concerning expectations, the reinvestment of part of the profits is not a necessary outcome of a decision to maintain a constant income stream. If the savings do result from the prospects of a high rate of profit, the rate charged on loans will be such as to limit the demand to the available supply, and the greater the demand for loans the greater, through the pressure of the price for loans, the effect on the price of part-substitute factors. This rise in costs, due mainly to the limited elasticity of the short-period supply of savings and the slow amortization of existing investments, sets a limit to the rate of expansion in the profit industries. and in so bringing about a gradual adjustment prevents the chaos that would otherwise arise as the result of statically established plans.

---

[1] In an article on 'The Equilibrium of the Firm', *Economic Journal*, 1934, Kaldor seeks to prove that in static perfect competition the size of the firm must continuously increase After objecting to making the equilibrium of the firm dependent on the equilibrium of the industry Kaldor proceeds, in effect, to ignore the industry. But it is wrong to suppose that supplies of factors are just used up, that prices of products are just such as to clear the market, and at the same time to assume that to the firm (any and every firm) the factors are in perfectly elastic supply and the demand for the product perfectly elastic at the equilibrium price.

It follows that the movement to a no-profit equilibrium is dependent, as in 'classical' theory. on both savings and investment being interest-elastic. One would expect, therefore, that corresponding to static perfect competition would be a 'classical' theory of employment. And so, it will be argued, there is.

3

So far it has been assumed that money is a medium of exchange only—i.e. that there is no deficiency of effective demand. Is this an arbitrary assumption, or is it implicit in the assumption of static expectations and rational action?

In his *General Theory* Lord Keynes argued that the rate of interest is not the 'price' which brings into equilibrium the demand for resources to invest with the readiness to abstain from present consumption, but the price which equilibrates the desire to hold wealth in the form of cash with the available quantity of cash.[1] As a consequence Lord Keynes's theory involved not only the replacement of the classical quantity equation relating money and money income by a liquidity equation relating money and the rate of interest, but also, since the monetary equation cannot determine both the price-level and the rate of interest, the introduction of a supply-of-labour equation in which, up to full employment, the supply of labour is perfectly elastic at the ruling money wage-rate.

According to Lord Keynes a necessary condition for the existence of a liquidity preference for money as a means of holding wealth is, given that the rate of interest is never negative, the existence of uncertainty as to the future rate of interest.[2] This means that static expectations and a liquidity preference for money as a means of holding wealth could not exist together. Thus the crucial question is whether the rate of interest could be other than positive. If savings are made for the sake of income yield the rate of interest must tend to equality with the representative man's rate of discounting future satisfactions, which, in the quasi-stationary circumstances we are considering, must be positive. It follows that if we include in the definition of rational action that the motive for saving is the monetary return to be gained therefrom, as did Ricardo,[3] the assumptions of static expectation and rational action[1] ensure that money exists as a medium of exchange only. That is to say, Say's Law, that there can be no deficiency of effective demand, holds.

---

[1] *The General Theory*, p. 167.

[2] Ibid., p. 168.

[3] See, for example, *Principles of Political Economy and Taxation*, pp. 68, 176, and *Letters to Malthus*, lxxvii.

However, if we follow Professor Pigou and introduce 'the present value of the direct amenity utility in the form of power, sense of security and so on, . . . which a typically constituted man expects to derive from *having* his marginal unit of present saving, as distinct from the utility which he expects the future income due to that unit to yield',[2] the rate of interest would equal the representative man's rate of discounting future satisfaction less an allowance for this amenity value, and thus might be negative. In such a case money could remain but a means of exchange as it can be held at no cost, only if people acted irrationally. In this case, without uncertainty as to future rates of interest, money could become a means of holding wealth. Does this mean that the pricing mechanism is incapable of ensuring full employment, even with static expectations and rational action (in Pigou's sense)? If net investment is nil but savings are positive, employment and real income must fall, but it does not follow that the fall in real income would, in reducing saving, bring about equilibrium. Competition for employment would progressively reduce money income without increasing employment. Professor Pigou has argued, however, that this very fall of prices would progressively diminish the amenity value of having savings, as distinct from deriving income from them, since the fall in prices would increase the real value of existing possessions.[3] Now if all prices are flexible (if not, the increasing burden of debts would offset any reduction in 'amenity value'), money is the only possession that could increase in value. The attainment of equilibrium would thus be dependent on an increase in the value of cash balances.

The objection to this argument is that the greater the necessary fall in the price-level the more impossible is it to abstract from expectations of a further fall in prices—i.e. the greater the clash between static expectations and rational action. It is true that Professor Pigou develops his argument with reference to a nil investment situation, but even then there is no proof that the unfavourable effects of non-static expectations *must* be overcome—the absence of any position of long-period equilibrium, if wages are flexible, is equally likely. The deduction of full employment from rational action and static expectations is firmly based only if a particular type of rationality is assumed—namely, the rationality of (classical) economic man. The classical economists' assumption that men save solely for the sake of the future income yield is an essential basis for their treatment of the rate of interest as the equilibrator of savings and investment.

---

[1] So long as these two assumptions are consistent, and their consistency implies a restriction of the problem to stationary conditions, or to departures from stationary conditions which can be described in terms of small movements along coherent curves.

[2] *Employment and Equilibrium*, p. 103.

[3] *Economic Journal*, 1943, pp. 349, 350.

There may in the 'classical system' be a less-than-full employment equilibrium, as, for example, when there are wage rigidities due to irrational action or monopolistic restriction. In so far as the irrational action or monopolistic restriction was manifested in a demand for a given money wage, the level of unemployment would be affected by a change in the quantity of money as this would change other prices relative to the price of labour. As this classical unemployment could exist only on the basis of a wage rigidity, it should not be confused with the Keynesian unemployment equilibrium. For in the former the unemployment could not exist without the wage rigidity—hence it is due, in perfect competition, to irrational action, while though the Keynesian unemployment *equilibrium* could not exist without wage rigidities, these rigidities cannot be treated as mere irrationalities since their absence would not lead to full employment, and their presence may be necessary for stability.[1]

In *The Keynesian Revolution* Dr. Klein complains that in accepting Leontief's assertion that the Keynesian system differs mainly from the classical system in not being homogeneous of order zero in prices,[2] Keynes practically admitted that unemployment could not exist without wage rigidities. 'Again as in the Treatise', Klein writes, 'Keynes did not really understand what he had written, and chose the wrong thing to publicize as his innovation.'[3] A reading of his reply to Leontief, however, does suggest that Keynes was not altogether dependent on Dr. Klein for an understanding of Keynesian theory. Mr. Keynes argued not only that there is abundant evidence to contradict the homogeneity postulate, but that Leontief's idea could be 'applied more fruitfully and with greater theoretical precision in connection with the part played by the quantity of money in determining the rate of interest'.[4] Now, as we have already seen, Keynes argued that a necessary condition for the existence of a liquidity preference for money as a means of holding wealth is the existence of uncertainty as to the future rate of interest. But the necessary conditions for the homogeneous system are two—static expectations and rational action (in Ricardo's sense).[5] It follows that the Keynesian system is basically incompatible with the homogeneity postulate. The part played by wage rigidities in the Keynesian system is an important one—the assumption that the supply of labour is infinitely elastic with respect to money wage-rates up to the point of full employment is, for example, necessary to determine the level of money prices—but the absence of wage rigidities does not guarantee full employment.

---

[1] See Keynes, *The General Theory*, p. 304, and Hicks, *Value and Capital*, pp. 269–70.

[2] *Quarterly Journal of Economics*, vol. li, pp. 192–3.

[3] Klein, op. cit., p. 83.

[4] *Quarterly Journal of Economics*, vol. li, p. 209; see also *The General Theory*, pp. 190–1.

[5] Cf. F. Modigliani, 'Liquidity Preference and the Theory of Interest and Money', *Econometrica*, vol. xii, pp. 67–70. Although in a homogeneous system the price-level is irrelevant the difficulty of injecting money without affecting the real part of the system is well known. See Wicksell, *Lectures* . . ., p. 159 et seq.; Keynes, op. cit.; Hicks, *Econometrica*, vol. v, no. 2, p. 151.

There is, nevertheless, one aspect of Keynes's treatment of expectations which rather obscures the significance of uncertainty. He wrote: 'an entrepreneur who has to reach a practical decision as to his scale of production, does not, of course, entertain a single undoubting expectation of what the sale-proceeds of a given output will be, but several hypothetical expectations held with varying degrees of probability and definiteness. By his expectation of proceeds I mean, therefore, that expectation of proceeds which, if it were held with certainty, would lead to the same behaviour as does the bundle of vague and more various possibilities which actually make up his state of expectation when he reaches his decision.'[1] This device is a necessary part of Keynes's 'static' method[2] and makes it possible to subsume considerations affecting the supply of risk bearing in the equations which determine the volume of investment —long-term expectations being treated as independent variables.

In Keynes's writings this device of a 'certainty equivalent' does not lead to any trouble in connexion with the homogeneous system, as his supply-of-labour equation is specifically non-homogeneous. In other hands the concept has caused trouble. Thus J. R. Hicks, after pointing out that people rarely have precise expectations at all, argues that the assumption of precise expectations can be made 'not altogether incompetent for dealing with a world in which risk is supremely important by taking as the representative expected price the most probable price ± an allowance for the uncertainty of the expectation'.[3] Clearly a unit elasticity of price-expectations derived in this way cannot in conjunction with 'rational action' (if in the context that term can have any 'classical' meaning) be a sufficient condition for a homogeneous system, and the use of this concept leads Hicks into error in the treatment of price flexibility. He first argues that where all elasticities of expectation are equal to unity, and the whole system of interest rates is taken as given, the system can be in equilibrium at any level of money prices.[4] This is immediately followed by an argument that a flexible rate of interest is a stabilizer in such a system,[5] which can be so only if the excess demand function for money is not a homogeneous function of the first degree in all the current and expected discounted prices, i.e. if excess demand functions for commodities are not homogeneous functions of degree zero. This contradiction can perhaps be explained by the fact that the effect of uncertainty in bringing a monetary influence into the determination of the rate of interest is obscured by the use of a 'certainty equivalent'.[1]

---

[1] Op. cit., p. 24, n. 3.

[2] Cf. Hicks, *Economic Journal*, 1936, pp. 238–53.

[3] *Value and Capital*, p. 126.

[4] Op. cit., pp. 254–7. For a criticism of this procedure see my note on 'Price Flexibility' in the *Economic Record*, Dec. 1946.

[5] Hicks, op. cit., pp. 258–9.

The part played by uncertainty is illustrated more clearly by the behaviour of savings and investment. In the Keynesian system there is nothing to ensure the equation of savings and investment at full employment income by the rate of interest. On the supply side, income is made the determining factor. On the demand side, the stress on psychological fluctuations in the marginal efficiency of capital, and on the sense of a normal rate in setting a limit to the fall in the rate of interest in the face of a reduced demand for saving, also reduce the rate of interest to a relatively unimportant place. Although Keynes did not drop interest from the demand-for-investment equation, there is good reason to argue that he overstressed its importance. For his so-called 'psychological law' that the marginal propensity to consume is less than unity is a statistical generalization, and the statistical evidence points also to the interest-inelasticity of investment. Furthermore, the maintenance both of a sense of a normal rate of interest and of the fluctuations in the marginal efficiency of capital which characterize the Trade Cycle are dependent on the existence of industrial progress.[2] But innovations will shorten the 'economic horizon' of large numbers of business men, and the shorter the horizon the smaller the influence of the rate of interest.[3]

The interest-inelasticity of savings and investment in the Keynesian system may be compared with their position in the classical system. In the former, the unimportance of interest in investment plans is explicable in terms of short horizons and the increased use of internal financing coupled with a failure to charge imputed interest. The first is due to uncertainty, the second to monopolistic competition as well as (where the failure to charge imputed interest is not directly irrational) to the use of rough-and-ready, 'rule-of-thumb', methods of procedure where close calculation would be impossible. In the conditions we examined in section 2, and in general for the conditions necessary for the classical procedure, there is, however, no reason to doubt the interest-elasticity of investment.

---

[1] See also M. Friedman's criticism of O. Lange's use of this concept in *American Economic Review*, Sept. 1946, pp. 627–30.

[2] Cf. Keynes, *The Treatise on Money*, vol. ii, pp. 95–6. There Keynes accepts Schumpeter's thesis that the major changes in investment are due to 'innovations made from time to time by the relatively small number of exceptionally energetic business men'. Keynes adds: 'the pace at which the innovating entrepreneurs will be able to carry their projects into execution at a cost in interest which is not deterrent to them, will depend on the degree of complaisance of those responsible for the banking system'. Now as the phenomenon of credit creation is, in the end, explicable only in terms of its use to enable or facilitate the introduction of new production functions, and as this mechanism involves the contraction of debts in money terms, changes in MV are intimately related to changes in relative prices. The proof of the 'non-homogeneity' of the Keynesian system can be made quite direct. Cf. p. 135, n. 1, and p. 136, n. 5, above.

[3] Cf. Hicks, op. cit., pp. 225, 226.

The argument for the interest-inelasticity of saving is as follows. There is no significant correlation between consumption and interest rates when the correlation between savings and income is taken into account.[1] Although some savings respond positively to variations in the interest rate (savings for wealth accumulation), other savings respond negatively (savings for annuities), leaving the total effect in doubt as to sign. Savings, in fact, are regulated largely by habits and considerations of economic security and they have little to do with the rate of interest.

The very form of this argument is such as to suggest that differences between the classical and Keynesian conditions are much the same in the case of savings as in the case of investment. The supply of savings at any time depends both on the rate of interest and on real income. As it is clear from the previous argument that in conditions where there are static expectations and rational action (in the Ricardian sense), short-period real income can be treated as constant, for the classical system we are left not only with saving as a function of the rate of interest but with 'savings for wealth accumulation'.[2] When we proceed to the Keynesian system the reaction of savings to the rate of interest is vastly different. First, there is no limitation of saving to saving for the income yield as is evident from the use of the annuities example.[3] Second, where saving 'by habit' is not in conflict with the marginal principle, the uncertainty concerning prospective income streams places a premium on conventional or rule-of-thumb principles of saving which automatically gives an initial passive relation of consumption to income.[4] Finally, reasoning on the lines of diminishing marginal utility is strictly applicable to the relation between consumption and investment only if they stand in a competitive relationship, which they do not do when there are unemployed resources.

It follows that we are on safe ground in dismissing Dr. Klein's complaint that in admitting his system to be 'non-homogeneous' Keynes practically admitted that wage flexibility could cure unemployment. The non-homogeneity of Keynes's system cannot be 'cured' by assuming away wage rigidities. In the classical system unemployment could be treated as voluntary, precisely because the classical system can be correctly applied only to a situation which is quasi-stationary—to a situation which diverges from the stationary state only by growth which can be described in terms of movements along coherent curves.[1] In the Keynesian system the uncertainty which is an integral part of industrial progress is vital. The non-homogeneity of his system could be cured not by imposing wage flexibility but only by abolishing something that is essential to capitalism—viz. entrepreneurship.[2]

---

[1] Klein, op. cit., p. 60.

[2] It was argued at the end of section 2 that savings and investment must be interest-elastic if a no-profit equilibrium is to be attained. As a positive reaction of saving to an increase in the rate of interest is a necessary condition of free entry and not of full employment, the reasoning in the text above is not circular. In the case considered in section 2, little can be said concerning the effect of real income on saving as the change in utility functions renders impossible comparison between real income in the initial situation and real income during the adjustment. But there is nothing in the nature of the change in demand to induce new savings apart from the effect on the rate of interest.

[3] Savings for annuities, implying a monetary constant, are inconsistent with a homogeneous system—a further illustration of the close confines of such a system.

[4] Kuznets has shown that for the U.S. the ratio of saving to income has been stable since 1870, with some declining tendency after 1900. But during the ups and downs of the business cycle the savings ratio deviates considerably from its long-run level, and does so in a systematic fashion—saving lags behind its secular ratio during the down-swing, and vice versa. See F. Modigliani, *Social Research*, vol. xiv, no. 4.

4

It remains to consider explicitly whether there is any meaning in Keynes's attempt to prove that there will generally be an equilibrium with unemployment in perfect competition. We have seen that Keynes's system implicitly depends on the existence of dynamic competition. But in dynamic competition, as Schumpeter has shown, 'the modus operandi of competition is *different in kind* from that in static competition',[3] and, it will be argued below, there is no such tendency to that equality between prices and costs which is usually implied by the assumption of perfect competition. The fact that it has been possible for Keynesians to assume perfect competition is due in part to carelessness in the treatment of free entry, and in part to certain ambiguities in Keynes's own treatment.

In the Preface to *The General Theory* Keynes wrote:

> 'This book . . . has evolved into what is primarily a study of the forces which determine changes in the scale of output and employment as a whole; . . . Our method of analysing the economic behaviour of the present under the influence of changing ideas about the future is one which depends on the interaction of supply and demand, and is in this way linked up with our fundamental theory of value.'

Later he wrote:

> 'My next difference from the traditional theory concerns its apparent conviction that there is no necessity to work out a theory of demand and supply of output as a whole. Will a fluctuation in investment, for reasons just described,[4] have any effect on the demand for output as a whole ? . . . Traditional theory makes no answer at all, never having given the matter a single thought.'[5]

As a result of this conception of the problem Keynes has both explicitly and implicitly stressed the continuity of his General Theory with the classical value theory. Now, of course, all value theory has not been based on the identity of price-expectations and current prices, and in so far as value theory departed from static perfect competition but was still content to take over the conclusions of the Ricardians concerning employment, Keynes could legitimately argue that he gave an answer to a missing question. But in so far as Keynes depends on value theory for certain assumptions about an achieved equilibrium between prices and costs, which after all is a 'normal' or long-period relationship, neither the continuity with value theory nor the consistency of Keynes's method is so obvious.

---

[1] See Schumpeter, 'The Instability of Capitalism', *Economic Journal*, 1928.

[2] See Schumpeter, *Capitalism, Socialism, and Democracy*, chapters 11 and 12.

[3] Schumpeter, op. cit., chapters 7 and 8.

[4] The reasons are the dependence of investment decisions 'on the two sets of judgement about the future, neither of which rests on an adequate or secure foundation'.

[5] *Quarterly Journal of Economics*, Feb. 1937, p. 219.

From the assumption of perfect competition Keynes deduces the equality of price and marginal cost, and (in a closed economy) an adjustment of prices such that the money wage is equated to the value of the marginal product of labour. From the additional assumptions of given techniques and quantities of equipment he deduces the existence of diminishing returns and a fall in real wages with an increase of employment. This approach suggests that this theory moves within the familiar bounds of static perfect competition, even though it is an approach quite unjustified by his stress on uncertainty, and on user cost as a link between the present and future.[1] For unless expectations are uniform there will be differences in the marginal costs of firms situated similarly in all respects but for expectations, and 'the special mark of the fully dynamic problem is the difference between different individuals'.[2] And although the assumption of given technique, of the quality and quantity of equipment, is needed in his theory to permit employment to be treated as approximately proportional to income (output), it is useless for an explanation of the actual course of real wages when employment increases, since such an explanation involves consideration of changing technique and imperfect competition and is not explicable in terms of a movement along given cost curves. 'Diminishing returns' deduced from the assumption of perfect competition equilibrium has no valid automatic place in his theory.

Second, Keynes's use of the term 'marginal efficiency of capital' allied with the doctrine that the rate of investment will be pushed to the point on the investment demand-schedule where the marginal efficiency of capital in general is equal to the rate of interest, involves a suggestion of marginal productivity theory.[3] Although he stated that 'knowledge of the factors which govern the yield of an investment some years hence is usually very slight and often negligible', and that 'those who seriously attempt to make any such estimate are often so much in the minority that their behaviour does not govern the market',[1] he still placed considerable stress on the interest-elasticity of investment as in marginal productivity theory. As a result this basic 'model' and the development of his argument are often in conflict.

---

[1] *The General Theory*, pp. 66–73.

[2] Lindahl, *Studies in the Theory of Money and Capital*, p. 349.

[3] See also his explanation of investment in the U.S. in 1929 on p. 323 of *The General Theory*. Cf. also p. 178, and the comments on 'financial prudence' at p. 100.

Third, in his treatment of saving Keynes renamed 'time preference' as 'the propensity to consume'.[2] Yet when arguing that the propensity to consume is a stable function of income, reliance is placed for its short-period behaviour on quite different considerations. 'In the case of the so-called cyclical fluctuations of employment . . . habits, as distinct from more permanent psychological propensities, are not given enough time to adapt themselves to changed objective circumstances.'[3] It is clear that his psychological laws, whether of consumption or liquidity preference, have nothing in common with, for example, the psychological assumptions at the back of the law of individual demand.

Finally we may refer to his use of the concepts of supplies and demands as identities rather than as schedules. By treating supplies and demands as necessarily equal—as but two sides of the same transaction—he was able to treat what is not sold as a reserve demand dependent on (given) expectations. Thus by simply adding to the usual static assumptions of given tastes and resources certain assumptions about anticipations he was able to treat price as determined by current supplies and demands and utilize familiar equilibrium analysis. Furthermore, by the assumption that short-term expectations may largely be ignored[4] (since his distinction between short- and long-term expectation is the basis of his distinction between consumption and investment[5]) he was able to utilize equilibrium analysis even in a position of disequilibrium brought about by a change in long-term expectations (independent variables).

The benefits and dangers of this method have often been discussed, and here it is only necessary to point to the bearing of his procedure on value theory. In terms of the factors necessary to explain the uncertainty postulated by Keynes, there will not be that tendency to an equilibrium between prices and costs which is the substance of static theory. Confusion results from assuming the existence of such a continuity from a mere method of analysis. As the problems of dynamic and static theory are different in essence, and as the concept of an equilibrium in perfect competition does carry with it the implication of an equality between prices and costs, the attempt to demonstrate the compatibility of involuntary unemployment and perfect competition equilibrium is meaningless.

---

[1] *The General Theory*, pp. 149, 150.

[2] Ibid., p. 166.

[3] Cf. p. 138 above.

[4] Op. cit., p. 50.

[5] Ibid., pp. 46–7.

5

It was argued at the outset that the developments in the theories of effective demand and the firm had not led in the same direction, and we are now in a position to consider some of the implications of the Keynesian system for the theory of perfect competition as expounded in the *Elements of Economics*.

First, if it is static perfect competition which is expounded in the elementary courses, there is no justification for combining it with a 'Keynesian' explanation of unemployment. Where it is found desirable to deal with both, error is avoided only by maintaining a division between value and monetary theory.[1] That is not to say that some explanation cannot be given of the lines on which division between them is reduced in more advanced courses, but that is a different procedure from making the theory of employment grow out of the theory of individual choice.

Second, where because of non-static expectations it is necessary to take account of user cost as a part of marginal cost, costs cannot be treated as a simple function of output or the firm's output as determined by the equation of marginal cost and marginal revenue. For user cost is affected by expectations,[2] and with an expectation of a change in price, marginal cost is not independent of price. That partial equilibrium analysis which is implicit in an exposition by plane geometry is thus inconsistent with non-static expectations.

Third, in introducing user cost, the concept of cost ceases to have that objective nature which is implicit in optimum-allocation theory. For divergent expectations about an uncertain future will lead to variations in marginal cost between firms which, other than in their expectations, are situated similarly in all respects. In fact the doctrine that competition brings about the greatest satisfaction of consumers' desires at the least cost is a deduction from static perfect competition. But the stationary or quasi-stationary state does not provide a maximum maximorum, yet in moving to higher levels of real income under the impact of entrepreneurial action, the pricing system loses in efficiency and cannot even guarantee a tendency to full employment. The concept of the consumer as king must give way before the concept of the entrepreneur as the initiator of a process of discontinuous change which does not lend itself to a description in terms of a theory of equilibrium.[3]

Fourth, if saving is not interest-elastic there can no longer be a *general* presumption in favour of either the possibility or the practicability of free entry. Where there is full employment, but savings are not interest-elastic, only forced savings could bring an increase in the supply of investment. But such an inflationary process could not ensure an equilibrium. With unemployment, a greater demand for investment could, without forced saving, bring an increase in supply, either through a ready reduction of liquidity on the part of the banks, or by inducing dishoarding through a rise in the price of loanable funds.[1] But in conditions which explain an unemployment 'equilibrium', in the sense that price flexibility does not induce an adjustment in the direction of full employment,[2] there would be no such tendency for a fine calculation between prices and costs as is usually supposed when dealing with free entry in competitive conditions. Where enterprise is but 'little more than an expedition to the South Pole based on an exact calculation of benefit to come',[3] there is no justification for assuming that the ability to get finance for investment must (rationally) lead to its utilization.

---

[1] Cf. Keynes, op. cit., pp. 293–4.

[2] See Keynes, op. cit., chapter 6, Appendix, and A. C. Neal, 'Marginal Cost and the Dynamic Equilibrium of the Firm', *Journal of Political Economy*, Feb. 1942.

[3] See Schumpeter, op. cit., Part 2.

In concluding, we may take the opportunity of pointing to one of the important practical dangers of confusing static and dynamic competition. Recently there has been a considerable literature in favour of equality between price and marginal cost as *the* criterion for managers of State enterprises. But in dynamic competition, not only does the concept of marginal cost become rather ambiguous, but since the pricing system cannot be relied on to co-ordinate supply and demand as a whole, no enterprise in isolation would, in calculating costs from present and expected prices, be implicitly counting the cost to the community. As already said, marginal cost is an objective magnitude only in stationary conditions. Yet presumably it was not the conscious intention of all the exponents of the marginal-cost principle to imply that when we come to State enterprises the realistic assumption is that of the stationary state.

QUEEN'S UNIVERSITY,
BELFAST.

---

[1] The finance of investment must precede the supply of saving out of a greater output.

[2] Cf. Pigou, *Employment and Equilibrium*, chapter 4.

[3] Keynes, op. cit., p. 162.

THE FUNCTION OF EXCHANGE RATES

By R. G. HAWTREY

SIR HUBERT HENDERSON's article in the January issue of *Oxford Economic Papers* deals with the problem of monetary stabilization, and particularly with the mutual relations of the price level and wage level at home and abroad, employment, rates of exchange, and the balance of payments.

He contests 'the idea that variations in foreign exchange rates can serve as an efficient regulator of a freely working international economic system' (p. 1). To justify it, a depreciation of a country's exchange rates must be assumed to effect a lasting improvement in its balance of payments (p. 5).

No doubt 'it is normally helpful in the long run to a country's balance of payments that its currency should be undervalued rather than overvalued'. But 'the undervaluation which has to be established is not a simple entity, but a tripartite relation, in which rates of exchange are only one factor, the other two being the level of costs and prices in the country concerned, and the level of prices in the outside world' (p. 6).

When world prices are falling, the effect of a depreciation of exchange rates 'can reasonably be expected to be a beneficial one, from the standpoint of the balance of payments'. It will not 'raise internal costs and prices, except possibly by way of recovery from a very recent fall'.

On the other hand, 'when world prices are rising, or even when they are more or less steady, . . . lowering of exchange rates would cause internal prices to rise' (p. 7).

'The extent of an inflationary movement, once it has been set going, cannot be forecast. The process is apt to be self-perpetuating, and indeed to gather momentum, unless and until it is checked by some obstacle or counter-force.' If growing inflation is allowed to be the occasion of further exchange depreciation, there will be no end to it.

The alternative is 'to call in aid whatever instruments may be most effective for forcing prices down, including perhaps the restriction of credit, These are the instruments which were brought into play more or less automatically under the gold standard.' Then, 'if it may be necessary, after exchange rates have been lowered, to use the traditional deflationary instruments to establish a certain degree of under-valuation, why not use them in the first instance to establish the under-valuation at the original parities?' (p. 8).

Sir Hubert Henderson generalizes his problem: of the three factors, exchange rates, internal prices, and world prices, one, world prices, is outside the control of national policy. 'Therefore it must be possible to cause one of the other two factors to move in whichever direction is required, while the third is held firmly at a fixed point, or at least is effectively prevented from changing in equal proportion. It is possible to cause internal prices to fall . . . by the use of deflationary monetary and financial instruments, while the exchange rates are held firmly at fixed parities, though there are limits to the extent to which this process can be carried out in practice. This was the *modus operandi* of the gold standard' (p. 9).

But, he contends, it is not possible to let exchange rates vary and keep internal prices stable. 'For the deflationary or counter-inflationary instruments are not precision instruments. If applied with sufficient force and for a sufficient length of time, they are capable of reversing an inflationary trend; they are incapable of the delicate work of allowing prices to rise so far and no farther. Still less could they hold prices and money incomes at a constant level if exchange rates vary.'

Then, 'if the criterion for regulating purchasing-power must be shifted in practice from the fair-seeming but deceptive one of the state of employment . . . might not the most satisfactory alternative in some conditions be the time-honoured criterion of the size of the monetary reserves, suitably adapted to modern conditions?' (p. 13). The suggestion is that, 'if and when our monetary reserves fall below a defined figure, chosen as the minimum compatible with safety, a substantial reduction of the note issue must follow automatically. . . . A reduction of the note issue would of course carry with it higher short-term interest rates and a restricted volume of bank credit.' Sir Hubert Henderson regards this as 'a clumsy method of curtailing aggregate demand', and 'in ordinary circumstances' would prefer others, 'such as an over-balanced budget and the limitation of long-term investment work' (p. 14). But when reserves are 'falling to the danger point . . . the method of credit restriction, clumsy though it may be, is one of proved if limited effectiveness; it would be essential to fall back on it as a weapon of last resort, and prudent to provide in advance that this should be done in time'.

That credit restriction through high short-term interest rates is in the last resort indispensable as a means of checking inflation is an essential principle of monetary control. An 'over-balanced' budget may suffice, if the surplus is applied to extinguish bank credit. But something further is needed to guard against bank credit being re-created through new bank advances to traders and individuals. The 'limitation of long-term investment work' may help, provided the reduction of capital outlay by public authorities does not simply release resources from the investment market for increased capital outlay by private enterprise.

Sir Hubert Henderson's argument is that a policy of maintaining full employment by stabilizing the price level or the flow of money is impracticable, because the only really effective instrument for checking an undue expansion is 'clumsy'.

That is true enough of an over-balanced budget or a restriction of capital outlay. Capital outlay cannot be quickly retarded; budgets are made once a year (or exceptionally twice). These expedients are slow and not sure. They would take many months to produce any effect, and even then the effect might be brought to naught by an expansion of bank advances.

But bank rate is on quite a different footing. Short-term borrowing reacts promptly to the short-term rate of interest. Most of the spending for which traders or consumers borrow for short periods can be postponed or curtailed with but little inconvenience.

When currency is supplied to a community through the short-term lending of a central bank of issue, that bank can cut off the supply altogether by simply refusing to lend. A century ago experience had shown that a flat refusal to lend was too violent a measure, and so was devised the milder method of lending at a charge high enough to have a deterrent effect on borrowers. Bank rate could be adjusted by fractional advances or reductions at a time.

It was practical experience of this method that led the Macmillan Committee to call bank rate 'a most delicate and beautiful instrument for the management of a monetary system'. Bank rate, it is true, was not used to maintain a stable flow of money, but to adjust the flow of money to a fixed gold parity, or in other words to fixed rates of exchange. But its power of delicate adjustment was no less tested by that objective than it would have been by a stabilization of the flow of money or of the price level. Small changes in bank rate were perpetually being made in order to correct undesirable expansions or contractions. The actual level of bank rate at any moment was that judged suitable to the conditions at the moment. If the conditions changed, bank rate could be changed at a week's notice (if need be, at an hour's).

The criterion relied on was the state of the gold reserve. But a monetary expansion caused no loss of gold if there was an equal monetary expansion in progress in the other gold standard countries. The system permitted of a world-wide monetary expansion, in which they all kept pace with one another.

When hand-to-hand currency could only be supplied by gold coin or by notes backed by gold, a world-wide expansion would result in a drain of gold into circulation, and would reach its limit when the drain reduced gold reserves below the prudent level.

The weak point of the system was that this conjuncture gave too tardy a warning. Steps were only taken to correct a world-wide monetary expansion at a stage when it had been gaining momentum for several years, and a severe application of dear money was needed to reverse it.

The expansion had not merely to be stopped but to be *reversed,* because the flow of money had been allowed to grow beyond what the available stock of monetary gold would support.

At first movements of expansion had culminated in financial crises with an 8 or 10 per cent. bank rate. After 1873 expansions were kept within more moderate bounds, and 6 or 7 per cent. sufficed to check them. In that period, up to 1914, a vitally important feature of bank-rate policy was the *prompt reduction* of bank rate, as soon as dear money was clearly seen to have done its work, and the corner had been turned.

That it has become possible nowadays to call bank rate a clumsy instrument is to be explained, I think, by the fact that since 1918 it has been clumsily used. The 7 per cent. bank rate imposed in April 1920 was maintained for many months after the inflationary movement had been quite evidently reversed and after growing unemployment showed that a relaxation of credit was what was needed. In 1929 the inflationary tendency in America had been successfully kept in check, and a palpable recession in industry had begun several months before the stock-market crisis of October and November, but the transition from dear to cheap money in both New York and London was fatally prolonged for five months. In 1931 the relief gained by Great Britain from the abandonment of the gold standard was disastrously impaired by the imposition of a 6 per cent. bank rate for five months.

Sir Hubert Henderson qualifies the power of bank rate to reverse an inflationary trend with the condition, 'if applied with sufficient force and for a sufficient length of time'. But in all these cases the fault was that it was applied for much more than a sufficient length of time. What is to be inferred from past experience is not that bank rate has to work by bludgeon blows, and cannot produce any intermediate result between a big deflation and none at all, but that a judicious adjustment of the time and the rate can produce a nicely calculated effect.

If that is so, bank rate can be used as well for maintaining an equable flow of money and price level as for preserving a gold parity and fixed rates of exchange.

Sir Hubert Henderson offers us no argument to the contrary. Nor, I think, does anyone else. The odium attaching to bank rate is due not to any reasoned conclusions, but simply to the plain fact that it has been so used since 1918 as to cause those devastating depressions by which the world has been beset. Curiously enough, the same school of thought which rejects the use of bank rate on the ground that it is unduly deflationary insists that depressions cannot be explained as attributable to deflation but must be caused by some deep-seated maladjustment of economic forces.

Insufficient account is taken of the principle of the inherent instability of credit, that any departure from monetary equilibrium, whether expansion or contraction, tends to start a vicious circle of inflation or deflation and to grow with cumulative force. A slight monetary disturbance may cause a disproportionately great effect. To preserve an equable flow of money and avoid both inflation and deflation, *early* action is necessary, and early action requires unceasing vigilance to detect the beginning of a movement.

Perhaps all the methods of monetary regulation can be called 'clumsy', relatively to the nicely calculated application which is required of them. But the term is only appropriate in this relative sense. Credit restriction is the instrument in the hands of the agency which actually supplies currency, the central bank of issue. It is that agency which has the power of cutting off the supply, and it can exercise just so much of this power as circumstances require by a suitable adjustment of bank rate.

Sir Hubert Henderson at any rate does advocate the use of credit regulation as an instrument of monetary policy. He would resort to it when other methods of regulating the flow of money have failed to prevent an excessive loss of monetary reserves.

A loss of reserves is the sign of an adverse balance of payments, and his thesis is that the adverse balance should be corrected by credit restriction and not by depreciation of rates of exchange. Of 'the three factors, exchange rates, internal prices, and world prices' he would choose to fix the exchange rates, and to allow internal prices to vary, abandoning the hope of adapting monetary policy to maintain full employment.

But if monetary regulation, in spite of its imperfections, can maintain fixed rates of exchange, why can it not maintain a stable price level? Fixed rates of exchange tie the internal price level to the world price level. They could not be maintained 'if in the future severe world depression and deflation should recur'; and, 'for different and more complex reasons, a continuing inflation of world prices would also prove incompatible with fixed exchange rates' (p. 15). But it is only 'in times of extreme disorder' (p. 17) that they would have to be abandoned. And, subject to that exception, Sir Hubert Henderson would base monetary policy on fixed rates, so that any fluctuation in world prices short of 'extreme disorder' would be reflected in internal prices.

His case against varying rates of exchange is based on the reaction of the external price level on the internal through the rates of exchange (pp. 6–9). This reaction occurs no less when the rates of exchange are kept unchanged and the world price level changes.

The world has had long experience of the maintenance of rates of exchange at fixed parities. When credit regulation is used to correct any disturbance of the balance of payments, the purpose is to bring the internal price level into conformity with the external. To employ credit regulation to maintain a stable price level (or preferably an equable *flow of money*, through incomes, outlays, and disposals) would in one respect be more difficult: there would be no such simple indicator of a departure from the norm as is furnished by the state of monetary reserves. There may be doubt whether the circumstances of the moment require measures of expansion or contraction. So long as there really is doubt, there is probably no harm in waiting, provided action is taken as soon as the doubt is removed.

Sir Hubert Henderson is directing attention to the case of a single country confronted with the task of regulating the flow of its own money, while the world price level 'lies outside the control of national policy' (p. 6).

When he says that 'a reasonable stability in the purchasing power of gold, if that is to remain the standard in terms of which exchange parities are defined, is an essential condition of international exchange-rate stability' (p. 15), what he has in mind is that in this respect 'extreme disorders' should be avoided.

'In practice,' he proceeds, 'the purchasing power of gold will depend mainly on the policies pursued by those countries which are untroubled by balance-of-payments anxieties and equipped with abundant gold reserves. Therefore a special responsibility for creating the conditions in which exchange-rate stability can be maintained rests upon these countries, outstanding among which is, of course, the United States. In the meantime, uncertainty upon this matter would suffice to forbid the idea of attaching to modern exchange parities the sacrosanctity which they possessed in the hey-day of the gold standard' (pp. 15–16).

What then? To say that there is 'uncertainty upon this matter' is an understatement. That monetary policy in the United States is *not* directed to maintaining 'a reasonable stability in the purchasing power' of the dollar, and therefore of gold, is only too certain.

Inflation has been visibly in progress in the United States ever since price controls were removed in 1946. From time to time there is talk of stopping it. At the present time there is a prospect of a restoration of price controls, but price controls are unlikely to be sufficient, so long as redundant money and easy credit are perpetually threatening to generate an excess demand.

The flow of money in the United States, it may be said, has already so expanded that the quantity of money (currency, *plus* bank credit) is no longer out of proportion to it. That is not certain, and, even if it were, it could not be inferred that inflation has run its course, so long as easy credit prevails. It is only too likely that at last, when a persistent recrudescence of inflation has become intolerable, there will once again be a disastrous deflation in the United States.

Gold is tied to the American dollar, so that stabilization of the purchasing power of gold means stabilization of the purchasing power of the dollar. Any idea of stabilizing either is remote from the views of the American Administration and Federal Reserve System.

Sir Hubert Henderson would ask of them no more than the avoidance of 'extreme disorders', but has nothing to say of the methods by which they should do so. When they decide that credit restriction is necessary, no doubt they will have to resort to the recognized methods of credit restriction. If they act with as heavy a hand as in 1920 the description of these methods as 'clumsy' will once again seem only too appropriate. But what criterion are they to look to in deciding *when* to act? Monetary reserves and rates of exchange do not help them. The money units of the rest of the world are tied to the dollar, and it is no use to seek a standard by tying the dollar to them.

The trouble to be guarded against is a loss of value in the dollar, that is to say, a rise in the American price level and wage level. And the real question at issue is how far the rise is to go before corrective measures are applied.

The wage level is the decisive factor. When the price level outstrips the wage level, it can be brought back without serious strain or dislocation. But when the high price level is reflected in a corresponding wage level, retreat becomes difficult. In the past it has only been possible to reduce the wage level by the pressure of unemployment—hence the odium attaching to deflation, and, through deflation, to the gold standard and bank rate.

Is a monetary expansion to be tolerated up to the point at which its reversal will just stop short of an 'extreme disorder'? If it is to be *reversed* at all, there must be some deflation and presumably some unemployment, a disorder falling short of the extreme. If it is not to be reversed, that means that the wage level must be held where it is.

Then why not adopt a monetary policy calculated to hold the wage level where it is from the beginning? Is that too delicate an operation for the 'clumsy' means of regulation? If so, the choice is only between the progressive degradation of money, euphemistically called an expansionist policy, and alternations of inflation and deflation.

In a sense the dilemma is a real one. For even if the apparatus of credit regulation is not clumsy at all, but is as 'delicate and beautiful' as the Macmillan Committee said, no one would claim that it can obliterate all fluctuations in the flow of money however minute and however short. The problem is to keep the monetary fluctuations within the narrowest limits practicable, and the solution is in *prompt action.* Once it becomes clear that the flow of money is expanding out of its due proportion to the wage level, restrictive measures should be applied without delay.

There is a real difficulty, in that there is no tangible index of this state of things. When the wage level actually begins to rise, action is already overdue. The directorate of the central bank must rely on their own judgement.

In practice the difficulty need not be serious. For the banking community is always fully aware of the state of business. When there is full employment, industry is always on the verge of becoming over-profitable and over-active. The sign of that happening is that increased orders are being received by producers. Increased orders will be reflected in increased demand for materials, which may or may not cause a rise in market prices. What is more fundamental is an increased demand for *labour*. That, combined with high profits, gives workpeople the opportunity to ask for higher wages.

The directors of the central bank ought to be in a position to detect this situation at least as soon as the workpeople themselves. Then is the moment to apply a touch of credit restriction, perhaps a rise of a half per cent. in bank rate. If that has no effect or insufficient effect, a further rise, and, if need be, successive rises, till the expansive tendency is checked.

That is the way bank rate has worked in the past, by trial and error. Some people fear that there can be nothing intermediate between a wholly ineffective rate and one excessively deflationary. But that is not what nineteenth-century experience teaches. To gradations in bank rate in those days correspond gradations in its restrictive effect. Since 1919 there has been only too much experience of an unduly deflationary rate. But that was either because dear money was applied when business was already depressed, as in 1925 and 1931, or because it was prolonged after it had accomplished its effect, as in 1920–1 and 1929–30.

When dear money has been resorted to, and an expansion successfully counteracted, prompt action once again is needed. To avoid a lapse into a serious deflation there should be an immediate relaxation of credit. This would conform to nineteenth-century practice, when the decisive application of dear money was regularly followed by a rapid transition to cheap money.

Stabilization of the flow of money by this method may be an elusive ideal, but it is the only criterion available to a country in the position of the United States, to which neither gold reserves nor rates of exchange can afford reliable guidance.

It is about British monetary policy, not American, that Sir Hubert

Henderson writes. His problem arises from the possible discrepancy between internal and external price levels, and the external means the American price level.

His recommendation is to maintain an unvarying dollar value of the pound, and to rely on the country's monetary reserves as the indicator in the regulation of credit to that end.

That would mean letting the pound remain among the crowd of money units which passively follow the dollar, and abandoning the leadership in credit policy and monetary policy to which London has hitherto aspired. It would also mean exposing the trade and industry of Great Britain and the sterling area to fluctuations in the wealth-value of their monetary units corresponding to the vagaries of the American dollar.

If we are not to submit to those vagaries, with the resulting alternation of inflation and deflation, the essential condition of escape is the release of the pound from its fixed rate of exchange on the dollar. Sir Hubert Henderson describes a depreciation of the money unit as beneficial, 'during a severe world depression, when prices in terms of gold and other currencies are falling sharply'. But why not beneficial when the fall of prices is not so sharp, and the depression not so severe? Why submit even to a moderate depression and moderate unemployment, if they can be avoided merely by an appropriate reduction of gold parity?

If the foreign-exchange value of the money unit is to be reduced when world prices are falling, in order to avoid deflation, then it ought likewise to be raised when world prices are rising, in order to avoid inflation. If rates of exchange were nicely adjusted to changes in the world price level in this way at suitably chosen intervals (possibly sometimes very short), then whatever mechanism was brought to bear to maintain the rates in force for the moment would in effect be tending to stabilize the internal value of the unit.

An unchanging price level is not a perfect solution. What is wanted is an equable flow of money. The difference is a complication into which I will not enter here. It is sufficient to say that, whatever the definition of undesirable variations in the real value of the money unit may be, if credit regulation is to be employed to prevent them, rates of exchange must be varied in such manner as to offset undesirable variations in the real value of foreign money units.

Sir Hubert Henderson especially deprecates a resort to exchange depreciation to counteract an adverse balance 'when world prices are rising, or even when they are more or less steady' (p. 7). Even when world prices are rising, the adverse balance may nevertheless be due to the money unit being over-valued. If so, a reduction of its foreign-exchange value may still be the appropriate remedy.

But no doubt Sir Hubert Henderson is here thinking of an adverse balance due to some *non-monetary* cause. There may be a deterioration of a country's exporting power, a loss of external sources of income, or an increase of external indebtedness. An adverse balance results, and, even with no change in the world price level, the existing rates of exchange no longer maintain equilibrium.

On this Sir Hubert Henderson is surely right. The power of a depreciation of the money unit to correct an adverse balance is *limited*: 'after a certain point the unfavourable effects of a further under-valuation on the terms of trade will outweigh the favourable effects on the volume of imports and exports' (p. 6). And 'the rise of internal costs and prices, following a lowering of exchange rates', may 'overshoot the mark and . . . result in an overvaluation' (p. 8).

Any rise in costs, so far as it goes, nullifies the advantage of depreciation. As soon as the depreciation begins to cause a rise in the wage-level, its usefulness has reached its limit.

And the limit is likely to be reached quickly. For the favourable effect of the depreciation arises from the increase of the external price level relative to the internal; the resulting increase in profit margins attracts producers to the production of exports and of goods which compete with imports; and these producers compete in the labour market for more hands. The workpeople have good grounds for asking higher wages when the employers are making excess profits and the cost of living has risen, and the employers are in a position to grant them.

It is no use pressing depreciation beyond the point at which the wage level begins to rise. What then is to be done? This is not a question of fluctuating rates of exchange. Given that the rates of exchange are already those which will have the most favourable effect upon the balance of payments, but that there is still an adverse balance, what is the remedy?

The adverse balance is the sign that the country is spending more than it is producing. It might produce more, but to that there is a limit, not only a physical limit, but a human limit. If measures can be taken for increasing production within those limits to a sufficient extent, that may be the best solution. But an increase in production is not to be had for the asking, and it cannot be assumed that the gap can be bridged in this way. The country is taking more imports, visible and invisible, than it can pay for. The immediate remedy is to restrict imports: visible, by import duties or prohibitions; invisible, by exchange control.

A restriction of imports raises the prices of the goods affected, and so tends to attract productive resources into the supply of these or of substitutes for them. Here once again, though on a more restricted scale, are excessive profit margins, a demand for labour, and a rise in the cost of living. If a rise in the wage level results, exporting industries will be put at a competitive disadvantage, and *further* restriction of imports will become necessary.

Yet for most countries an unlimited restriction of imports is not practicable; there is a hard core of imports deemed indispensable. If the adverse balance is of such magnitude that a restriction of imports up to this limit fails to overcome it, recourse must be had to measures of austerity. By rationing and restriction of supplies allowed to home markets, spending must be kept down, and productive power released for the production of exports.

In such circumstances, only too familiar to the present generation, the aim of monetary policy should be to maintain rates of exchange consistent with the greatest possible exporting power in terms of foreign money units *at the existing wage level.* The moment a depreciation of the country's money begins to cause a rise in the wage level, that means that it is going too far.

The adverse balance of Great Britain, its 'dollar problem', is of this character. Because it is not one of those countries 'which are untroubled by balance-of-payments anxieties and are equipped with abundant gold reserves', Sir Hubert Henderson would apparently regard Great Britain as disqualified to take the lead in the regulation of the world's monetary units. That leadership he thinks devolves on the United States.

Whether Great Britain can still lead remains to be seen. But at any rate it is not compelled to *follow,* if American leadership is misdirected. It can dissociate itself from the dollar, and gain a free hand in regulating the internal value of the pound. That would mean both regulating credit and adapting the rate of exchange at suitable intervals to variations in the wealth-value of the dollar.

The essence of this policy would be the regulation of credit in such a way as never to *influence* the wage level upwards or downwards, either by an inflationary widening of profit margins or by an incipient lapse into unemployment. The adjustments of rates of exchange would be ancillary to the regulation of credit, but would be in some degree independently determined by reference to the relation of export and import price levels to costs in the industries concerned with them.

The pound sterling so regulated would offer a really stable standard not only to British trade and industry but to any country desirous of linking its money unit to a standard free equally from deflation and inflation. It would be qualified to take the lead in a world monetary system.

Nor should the weakness of the British balance of payments stand in its way. British solvency may depend for a long time to come on the controls by which spending at home is restricted, and there is an obligation on the nation to impose on itself whatever degree of austerity is necessary to that end. If that is done, the weakness of the balance of payments will be corrected.

The money unit best qualified to lead the world will not be that best entrenched in monetary reserves and a favourable balance, but that which provides the most reliable standard of value. The monetary policy which will fit the pound sterling to take that place will be that which frees it from inflationary and deflationary influences.

At the present time the lead in monetary policy rests with the United States. Natural resources and productive power unrivalled and unimpaired by war, combined with 20,000 tons of gold, have secured that position. The immediate question is whether the pound should continue to follow the fluctuations of real value to which the dollar is subject. The indispensable condition of its being released is freedom so to vary the rate of exchange as to counter-balance those fluctuations.

THE ROYAL INSTITUTE OF INTERNATIONAL AFFAIRS,
CHATHAM HOUSE, LONDON.

A COMMENT BY SIR HUBERT HENDERSON

For the most part, Professor Hawtrey gives a very fair account of my argument; but the play which he makes of the words 'extreme disorder' might convey a misleading impression. I only used these words in my concluding sentence when I said that 'exchange-rate fixity, like other desirable things, may prove unviable in times of extreme disorder'. I was not referring to price movements there, and I did not intend to suggest, as Professor Hawtrey seems to infer, that only extreme changes in the level of gold prices could justify a particular country in altering its gold parity.

It may perhaps help to reduce misunderstanding and to define the issues more clearly if I enlarge a little on this matter. First there is the possibility that sterling may become over-valued. If this were to happen, solely because of inflation in the United States, and *after* the inflationary trend in Great Britain had been definitely arrested, I should agree that devaluation would be necessary. A country with a weak balance of payments cannot afford an over-valued currency; and, as one of the few who, from the beginning, and indeed before the event, criticized our return to gold in 1925 at the pre-1914 parity, I am not disposed to underrate the disadvantages, approximating in Great Britain to the impracticability, of correcting an appreciable over-valuation by internal deflation. In the light of what is now happening, it seems far more likely that sterling may become over-valued *before* our internal inflation is arrested. Dollar costs and prices might be falling, as they have been falling in recent months; or they might be rising, as they may conceivably rise again, though less rapidly than sterling costs and prices. But so long as inflation is at work in our own economy, devaluation would be a most dangerous remedy, threatening a cumulative process, like that we saw in France in 1948, of further inflation and further devaluation. The right policy is to try to avert this contingency by ending the inflationary trend at home; and the idea of devaluation should in my view be ruled out, until this has been definitely accomplished.

The chief practical difference between Professor Hawtrey and myself is about the circumstances which would justify not the depreciation but the appreciation of sterling. And here, of course, Professor Hawtrey differs far more sharply from nearly every other advocate of variable exchange-rates. The essential issue may be stated as follows: What should be done to sterling in conditions such as those which have prevailed during the last few years, assuming those conditions to include:

(*a*) a large deficit in the balance of payments; or, as to-day, the certainty that there would be one if import restrictions were removed;
(*b*) rising gold or dollar prices;
(*c*) a presumption, until recently at least, that sterling has been under-valued, rather than over-valued?

In such circumstances, most advocates of variable exchange-rates recommend, with differing degrees of reserve, that the gold or dollar parity of sterling should be lowered. Professor Hawtrey recommends, without any reserves at all, that it should be raised. I hold strongly that it should be kept unchanged. Most of my article, on its negative side, was directed against the devaluationists; and Professor Hawtrey would do wisely to recognize that if the pound were to be 'dissociated' from the dollar, it is their views that would almost certainly prevail. My reasons for dissenting from Professor Hawtrey's policy may be stated shortly as follows:

1. It would involve forcing a change against the current trend of supply and demand, or what certainly would be the trend in free exchange markets; and that is seldom a wise thing to do.
2. I do not believe that it would do much in conditions of general shortage, like

those of recent years, to check the rise in British costs and wage-rates. It might, therefore, lead to an over-valuation damaging to our export trade.

3. The relation between the levels of industrial costs in different countries cannot be calculated with sufficient accuracy. Changes in relative productivity are no less pertinent than changes in relative wage-rates.
4. Above all, there are important positive advantages in stability of exchange-rates. The one point in common between Professor Hawtrey and other advocates of variable exchange-rates is that they ignore these advantages altogether.

If a rapid rise of American prices were to continue indefinitely, I should agree that these considerations would be overborne. But, in view of our large adverse balance of payments with the dollar area, a dollar inflation would have to go much farther to justify putting the pound up than a dollar deflation would have to go to justify putting it down. Nothing, incidentally, seems more improbable now than a further large dollar inflation.

ALL SOULS COLLEGE,
OXFORD.

PROFITS IN BRITISH INDUSTRY FROM 1924 TO 1935

By RONALD HOPE

1. A Definition of Profits

1. No excuse is necessary for undertaking an investigation into the profits made in British industry between 1924 and 1935, since little enough is known of the operative principle in an economic system which was, in this period at least, primarily dependent upon the profit motive. The prospect of making a profit caused the wheels of industry to turn. But what exactly were these profits which it was hoped would accrue as the reward for enterprise? According to Professor Knight,[1] true profits are the balance of a business man's income after all relevant costs have been met and 'normal' wages of management and 'normal' interest upon capital invested have been deducted. Unfortunately, while perhaps satisfactory for the theorist, this definition is of little use to the statistician because profits in this sense cannot be measured.

2. Actual profits, in the sense of net returns to a given group of entrepreneurs, equal all receipts in respect of the year's activity minus all outgoings in respect of that activity minus an allowance for the deterioration of fixed capital invested in the business undertaken by them. Profits in this sense probably include some earnings for management and some interest on capital, as well as the reward for risk. It may be possible for the entrepreneurs concerned to make this estimate for any one business, but it is impossible for the economist to do so accurately from the figures of profit and loss published in the past. Even this concept may be criticized[2] on the grounds that all outgoings should be re-estimated at current prices and allowances for depreciation should be in terms of replacement cost, if true net returns are to be accurately assessed.

3. The Inland Revenue authorities have a definition of profits which is akin to, but not the same as, that suggested above. The Inland Revenue estimate of profits is that amount which the authorities agree to be the profit made by a company for any year for the purpose of estimating income-tax payments due to the Crown. It should equal the total receipts in respect of the year's activity, minus all costs incurred in that activity, and minus an allowance for wear and tear which probably approximates, over a period, to that which should be provided for the deterioration of fixed capital, although business men usually maintain that this allowance is inadequate and in many industries at the present time (1948) it

---

[1] F. H. Knight, *Risk, Uncertainty, and Profit* (Houghton Mifflin, U.S.A., 1921).

[2] Cf. K. Lacey, 'Profit Measurement and the Trade Cycle' (*Economic Journal*, December, 1947).

undoubtedly is. It includes all interest on capital and only a small part of it could be considered 'the reward for risk'.

4. In spite of its wide divergence from the theoretical concept of profit, the Inland Revenue estimate is a very useful one and nearest to that which we have in mind when considering profits in British industry. If the estimates made by the Inland Revenue authorities were made public, our task would consist of their analysis. With the single exception of the profits of brewing concerns, however, no public announcement has ever been made of the profits enjoyed by any industry, company, or group of companies. Our major task has been to endeavour in some degree to fill this gap in our statistical knowledge.

5. To this end, use has been made of the *shown profit,* which is that amount which the directors show to their shareholders as available for the general purposes of the proprietors of the company, including the payment of interest and dividends and the maintenance of reserves.

2. The Sample of Companies

6. The general aim of the inquiry was to investigate the way in which profits made by British companies varied from industry to industry and as a whole, over a period which was as long as possible and which covered at least one trade cycle. An investigation of this kind must be confined to a sample of companies, the sample consisting of those companies which published the requisite information over the whole of the requisite period. After companies not engaged in British industry proper—banking and financial companies, investment trusts, companies whose investments were mainly overseas, and companies registered outside the United Kingdom—had been excluded, 510 were found to yield satisfactory information. Although only 3·5 per cent. of the total number, these companies controlled just over 50 per cent. of the capital nominally invested in all public companies in 1935.

7. They were classified according to the industry to which they belonged, and in making this classification the broad outline of the Census of Production was followed. In some cases this involved a somewhat arbitrary choice. Companies show the greatest possible divergence from the statistician's ideal of one company, one product, and their activities often cover a whole range of industries. The criterion adopted was community of interest but, since the rule that a company should only go into one class was also adopted, this has meant that the boundaries between industries are sometimes obscured in borderline cases.

8. The shown profits of a past period, even where the companies are hand-picked, are subject to a number of limitations which are commented upon in the Appendix.[1] Some attempt has been made here to obtain

---

[1] pp. 179 et seqq.

comparable profit figures for each industrial sample by estimating taxes paid where these are not shown, but these estimates are subject to error. Whatever qualifications are made, however, there is little doubt that the information concerning fluctuations and trends in profits between 1924 and 1935 is, broadly speaking, correct. The figures also reveal incidental information of great interest. Most important is that relating to dividend policies, which merits fuller consideration than it can receive in a short article. The tentative conclusions which may be drawn concerning the general level of profits, too, are by no means unimportant.

3. Profit Fluctuations

9. Table I summarizes the profit experience of the major British industries between 1924 and 1935 as indicated by the published profit and loss accounts of the 510 companies examined. The figures given in the last column enable the reader to substitute cash for the indices shown and, in conjunction with the capital figures given in column 3, to examine separately the data for any particular industrial sample.

10. It has long been known that fluctuations in the capital goods industries are much greater than those in the consumer goods industries, but hitherto the data have been confined to output and employment figures. Table II shows fluctuations in profits between 1924 and 1935 in different industries measured from the mean profit for these years with trend excluded. The results conform closely to expectations, though fluctuations in this period may be abnormally high since it includes the year of the General Strike.

11. Fluctuations were most violent in the shipbuilding industry, being twice as great as those in the next worse group and eighteen times as violent as those in the food industry, the industry which suffered least from the effects of the trade cycle. The companies in the shipbuilding sample suffered heavy losses in each of the years from 1931 to 1933, and profits were negligible in 1926.

12. Next comes a group which includes coal, commercial motors, and miscellaneous engineering, with fluctuations only nine times as violent as those in food. Of these, the coal industry sample suffered losses in 1926 and 1928 and the miscellaneous engineering sample in 1933. It is interesting to note that the companies producing commercial motor vehicles (capital goods) suffered fluctuations in profits three times as great as those experienced by the vehicle accessories companies (partly durable consumer goods) and four times as great as those experienced by the non-commercial motor companies (wholly durable consumer goods).

13. Again there is a striking gap, with textile merchants, locomotive and wagon companies, iron and steel, and textile manufacturers being

TABLE I

Relative Shown Profits (including Interest) before Tax in the Major British Industries, 1924–35

(Average, 1924–35=100)

Industry (1)	*No. of companies* (2)	*Total share and loan capital nominally invested in 1935* (£ *mn.*) (3)	*Relative shown profits (including interest) before tax* *1924* (4)	*1925* (5)	*1926* (6)	*1927* (7)	*1928* (8)	*1929* (9)	*1930* (10)	*1931* (11)	*1932* (12)	*1933* (13)	*1934* (14)	*1935* (15)	*Average shown profits for each sample** (£ *mn.*) (16)
All industry . . .	510	1,995·43	111	100	85	115	115	**121**	102	89	*78*	84	95	105	130·20
Food	48	68·77	86	83	89	99	102	112	103	**114**	99	*98*	106	109	6·81
Drink	49	86·38	106	102	99	102	106	**116**	112	96	*76*	83	98	104	10·15
Tobacco . . .	9	91·24	86	91	99	102	105	110	**112**	104	96	*89*	100	106	20·14
Paper	22	51·32	89	92	93	104	111	**113**	102	*87*	92	100	106	111	5·15
Stores	21	26·92	106	116	111	113	115	**116**	98	85	*77*	83	86	94	2·65
Hotels and Restaurants .	19	29·38	107	104	100	105	107	**112**	108	85	*83*	93	98	98	2·22
Amusement . . .	6	7·65	62	77	92	114	110	132	**134**	97	91	*73*	103	115	0·66
Leather . . .	13	11·99	91	77	86	97	105	123	**126**	116	97	92	*88*	102	0·96
Clothing . . .	17	10·41	157	117	118	131	**147**	131	85	70	*47*	55	68	74	0·77
Building . . .	25	27·06	94	115	93	101	110	**111**	91	85	*75*	85	104	136	2·45
Non-commercial Motors .	8	10·50	87	109	89	83	81	**118**	95	94	*76*	112	134	122	1·36
Commercial Motors .	6	3·34	87	71	143	147	201	**215**	115	30	22	*9*	63	97	0·58
Vehicle Accessories . .	13	7·38	82	109	119	117	115	**125**	108	70	*60*	78	98	119	0·84
Electrical Engineering .	22	39·96	104	118	97	96	105	**114**	101	84	87	*71*	93	130	3·64
Chemicals . . .	19	17·22	92	83	79	99	**120**	112	96	89	*74*	92	124	140	1·50
General Engineering .	30	24·48	122	122	105	120	**130**	130	84	58	*39*	62	94	134	2·08
Miscellaneous Engineering	14	17·17	297	127	142	154	**191**	135	46	11	16	−*53*	38	96	0·52
Large Engineering . .	4	23·31	73	70	116	121	112	**122**	103	*82*	83	84	98	136	1·44
Iron and Steel . .	30	72·48	120	67	64	95	102	**122**	92	*59*	61	98	139	181	3·56
Coal	19	21·40	195	99	−2	15	−7	**136**	121	*83*	94	164	146	156	0·80
Railways . . .	4	1,102·40	118	112	53	122	118	**130**	108	95	*77*	82	91	94	34·58
Locomotives and Wagons	11	8·32	109	146	136	143	**156**	135	130	72	42	*11*	39	81	0·52
Shipping . . .	20	53·23	150	135	149	**161**	154	146	93	54	*26*	41	52	39	2·84
Docks and Wharves .	9	26·21	108	109	111	**121**	106	112	101	86	*76*	85	95	90	1·26
Shipbuilding . . .	10	27·94	163	153	6	119	134	**312**	141	−31	−72	−*75*	53	297	0·39
Textile Manufacturers .	35	108·67	178	151	107	**159**	151	105	42	*39*	54	76	65	73	7·46
Textile Merchants . .	18	11·25	236	200	104	130	**142**	126	63	57	*8*	36	46	52	0·61

---

* i.e. Average shown profits (including interest) before tax for each sample; for example, on an average, in each of the years between 1924 and 1935, the 510 companies which make up the complete sample showed a profit of £130·20 million.

subject to profit fluctuations only half as violent as these and some four and a half times as great as those in food.

14. In shipping, general engineering, railways, amusement, vehicle accessories, large engineering, and clothing we find fluctuations round about three times as great as those in the food industry. In chemicals, building, non-commercial motors, electrical engineering, and leather they are about twice as great, while in brewing, stores, docks and wharves, paper, hotels and restaurants, and tobacco, fluctuations are little more severe than those in the food industry itself.

TABLE II

The Degree of Fluctuation in Profits in the Major British Industries, 1924–35

Industry	*Standard deviation from trend (average 1924–35 = 100)*	*Industry*	*Standard deviation from trend (average 1924–35 = 100)*
All industry	12·5	Large Engineering	20·0
Shipbuilding	123·0	Clothing	19·9
Coal	60·5	Chemicals	16·5
Commercial Motors	59·7	Building	15·6
Miscellaneous Engineering	59·2	Non-commercial Motors	15·4
Textile Merchants	32·1	Electrical Engineering	15·0
Locomotives and Wagons	31·6	Leather	14·0
Iron and Steel	31·6	Drink	9·7
Textile Manufacturers	29·1	Stores	9·0
Shipping	24·3	Docks and Wharves	8·6
General Engineering	22·7	Paper	8·0
Railways	20·8	Hotels and Restaurants	7·5
Amusement	20·7	Tobacco	7·4
Vehicle Accessories	20·6	Food	6·8

15. In general, fluctuations become markedly less severe as we pass from the capital goods and export industries to the durable consumer goods industries, and from these to the non-durable consumer goods industries. The only exceptions to the rule appear to be docks and wharves—a service industry in which fluctuations were mild—and amusement. The latter is a small sample, consisting of a few theatres and cinemas which appear to have been subject to particularly acute competition in the early part of the period. Fluctuations in capital goods industries were, without exception, over two and a half times as great, while in durable consumer goods industries they were nearly twice as great, as those in non-durable consumer goods industries.

16. The small fluctuations in profits as a whole are due in part to the importance of the railway companies in the sample, but also indicate that in the added figures there is a good deal of cancelling out. That is to say, fluctuations in different industries are by no means always in the same direction. Most industries and most companies were earning smaller profits in the early thirties than they were in the late twenties, but it would, of course, be a mistake to suppose that in any given year all industries were either prosperous or depressed.

4. Peaks and Troughs

17. For trade-cycle theory it is important to establish which industries feel the onset of depression first and in which industries appear the first signs of recovery.

18. *A priori*, annual profit figures would seem scarcely sufficient for this purpose because effects are likely to be transmitted more quickly—within a few months—from one industry to another. Quarterly figures might be more useful, though a study of the quarterly figures of dividend payments, such as those published by *The Economist,* would probably prove of very limited value, since company savings usually absorb the first shock of a decline in earnings. Even were quarterly figures of earnings easily available, it is unlikely that they would reflect changes as quickly as figures for employment and output. One would suppose that the first repercussions of an oncoming depression would be a diminution in output and growing unemployment as the result of the cancellatian of orders, and only in a few prosperous industries is there any indication that employment and output may have been restricted because profits had fallen first. Even in these cases the evidence is extremely slight.

19. Annual profit figures, on the other hand, have the advantage that they eliminate chance and seasonal fluctuations and reveal, with few exceptions, more or less clearly defined peak and trough years. These years are indicated in Table I.

20. Profits showed signs of declining in shipping, docks and wharves, and textiles as early as 1927–8, and there were further signs of a decline in profits in a number of capital goods industries in the year which followed. The slump may have had its origin overseas, though the evidence is by no means conclusive. It is fairly clear, however, that the capital goods industries (with the exception of the motor industries and electrical engineering) were affected first. Peak profits were earned in no consumer goods industry, with the exception of the durable goods industry of clothing for which there was an important overseas market, prior to 1929. The highest profits in leather, tobacco, and amusement were earned as late as 1930, though the increases between 1929 and 1930 were slight. In the food industry profits fell from 1929 to 1930 but reached their peak for the period as a whole as late as 1931.

21. In most cases employment figures[1] gave warning of the slump at least as early as, and in many cases well before, profit figures. An exception to this rule, however, is to be found in the prosperous and expanding industries. In the chemicals industry sample profits reached their peak in 1928, employment in 1929. In electrical engineering the years were, respectively, 1929 and 1930. In the distributive trades, hotels and restaurants, and amusement, employment was increasing throughout the period, though profits among the sample companies showed a definite cyclical movement.

22. It is much more difficult to say anything very definite about the recovery period. In some of the basic industries—coal, iron and steel, railways, and shipping—there seems to have been a profit revival before the numbers employed increased. In the prosperous industries—chemicals, electrical engineering, the motor industry, food, drink, and tobacco—employment turned the corner first.

23. In textiles and paper the tide turned for both profits and employment as early as 1931–2. (The textile industry was on all counts an early 'indicator'.) By 1932 many other industries were employing more people than they had in 1931, though 1932 was in general the worst year for profits. Shipping and shipbuilding, railways, locomotives and wagons were among the last industries to show signs of revival according to both profits and employment figures, though the profit revival was also late in showing itself in food, tobacco, amusement, commercial motors, electrical engineering, and leather.

5. Profit Trends

24. The trend of shown profits in different industries (making no allowances for increases or decreases of capital) over the period from 1924 to 1935 was estimated largely to exclude it from fluctuations, to which a measure had been given in Table II, and not because it is claimed that a straight line gives in most cases a good approximation to the course of profits. For many industries it would have been better to fit two lines, one for the period from 1924 to 1929 and another for the period from 1930 to 1935, since the huge fall in profits in 1930 was of the nature of a landslide, and profits rarely recovered to pre-slump levels in the ensuing five years. Profits for all industry represented by the sample were some 15 per cent. lower in the second half of the period than they were in the first half.

---

[1] The employment figures with which the profit figures have been compared cover the whole of the industries concerned and were lent to the writer by Mr. P. W. S. Andrews.

25. Nevertheless, the trends are of considerable interest. If we take the result for all industry as the norm, we find that all consumer goods industries with the exception of clothing are above the norm. That is to say, profits in each of the consumer goods industries apart from clothing were, on an average, either increasing throughout the period or falling at a rate which was slower than the rate at which profits were falling for all industry taken together. On the other hand, profits in each of the capital goods industries, with the exception of chemicals and some sections of the engineering (including electrical engineering) and motor industries, were, on an average, falling faster than they were in industry as a whole. The coal and iron and steel industries alone are excluded because the trend of profits in these industries is misleading. Owing to the coal crisis of the middle twenties, and partly in the case of iron and steel to the Iron and Steel Federation's successful attempt to increase profits in the last two years covered by the survey, the trend of profits in both these industries rises markedly.

26. Manufacturers of commercial motors were severely hit by the slump with the result that the trend of their profits was quite sharply downward. Manufacturers of cars for the consumer, on the other hand, weathered the early thirties fairly well and the trend of their profits was upward. Manufacturers of vehicle accessories, with profits falling slightly though not so much as those for all industry, came between the two.

27. Chemicals, food, amusement, leather, paper, tobacco, and building all showed rising trends. In the building industry the sharp rise in profits in 1934 and 1935 played a large part in the determination of the trend line.

28. Among the consumer goods industries the experience of chain stores and one or two multiple tailors is of particular interest. From 1929 onwards a parabolic curve would fit the data well. Profits for these companies did not decline at all during the slump and increased at a rapid rate in the last two or three years of the survey. The increase in profits in these two fields between 1930 and 1935 was unparalleled elsewhere.

29. Profits in the brewing industry and among hotels and restaurants fell, on an average, throughout the period, though not by so much as profits in all industry. Experience in brewing was, however, exceptional among non-durable consumer goods industries and probably reflects the decline in consumption in this period.

30. A falling trend in railways, locomotives and wagons, shipbuilding, shipping, docks and wharves, textiles, clothing, and miscellaneous engineering was accompanied by low nominal rates of profit in these industries in 1935.[1]

---

[1] See opposite, Table III.

6. Rates of Profit

31. Nothing is more difficult than estimating the rate of profit in industry, because there is no means of estimating actual capital. In Table III, column 2, shown profits (inclusive of interest and before tax) in 1935 have been divided by the nominal capital invested in each industrial sample in that year. Columns 3–5 show payments, gross of tax, made on different types of share, and here some check is possible, since *The Economist* publishes each year similar estimates for its sample of companies. Taking account of the differing size of the samples and the differing methods of classification, a rough conformity exists between these sets of figures for 1935, though one fact which may be significant comes to light.

TABLE III

Nominal Rates of Profit and Payment in 1935

Industry (1)	*Gross profit* % (2)	*Gross debenture interest* % (3)		*Gross payments on prior shares* % (4)		*Gross payments on ordinary shares* % (5)		*Net earnings for ordinary shares* % (6)
All industry . . .	6·1	4·2	(4·7)	4·3	(5·2)	7·7	(7·4)	7·2
Tobacco . . .	23·3	4·8	(4·5)	6·1	(6·1)	28·7	(19·4)	23·4
Commercial Motors . .	17·5	6·0	..	5·5	..	16·4	..	19·6
Non-commercial Motors .	15·7	5·9	..	6·6	..	20·4	..	25·6
Vehicle Accessories . .	13·6	5·7	(5·7)*	6·7	(5·1)*	11·3	(9·8)*	15·2
Drink	12·4	4·4	(4·4)	6·6	(5·4)	16·5	(13·9)	17·5
Chemicals . . .	12·3	4·5	..	7·5	..	15·3	..	14·1
Building . . .	12·2	4·9	(4·8)†	6·4	(5·9)†	14·2	(12·6)†	14·4
Electrical Engineering .	11·7	6·8	(5·1)‡	6·4	(5·6)‡	11·4	(9·5)‡	12·0
General Engineering .	11·6	8·6	..	6·5	..	9·1	..	10·8
Paper	11·0	4·9	(4·6)	6·9	(5·8)	14·4	(9·5)	14·3
Food	10·8	4·7	(4·6)§	6·6	(6·2)§	11·6	(10·9)§	12·0
Amusement . . .	9·9	5·0	..	7·0	..	8·0	..	11·8
Stores	9·2	5·3	(4·7)	6·1	(5·9)	12·3	(14·1)	11·9
Iron and Steel . .	8·9	5·1	..	6·2	..	7·1	..	9·4
Large Engineering . .	8·4	4·7	..	6·3	..	6·8	..	9·3
Leather	8·2	4·7	..	7·8	..	10·7	..	8·9
Hotels and Restaurants .	7·5	5·7	(4·6)	3·8	(5·9)	7·9	(7·2)	8·3
Coal	5·9	4·6	(4·7)‖	4·7	(4·4)‖	4·6	(4·5)‖	4·7
Clothing . . .	5·5	4·7	..	3·3	..	3·3	..	4·2
Locomotives and Wagons	5·1	3·9	..	2·9	..	4·9	..	4·7
Textile Manufacturers .	5·0	4·4	(5·1)¶	2·5	(3·7)¶	5·9	(5·2)¶	5·1
Docks and Wharves .	4·4	3·8	(4·1)	4·2	(3·8)	4·8	(1·2)	4·2
Shipbuilding . . .	4·1	5·3	..	..	..	3·9	..	5·7
Railways . . .	3·0	3·9	..	3·7	..	1·0	..	0·5
Miscellaneous Engineering	2·9	4·1	..	4·4	..	2·9	..	0·8
Textile Merchants . .	2·8	5·2	..	5·6	..	2·4	..	0·9
Shipping . . .	2·1	4·3	(4·6)	4·0	(4·3)	2·0	(1·5)	0·0

Figures in brackets taken from *The Economist*, 18 July 1936.

The Economist classification:
* Motor-cycle and aviation. † Building materials. ‡ Electrical equipment.
§ Food, confectionery, and drink (non-alcoholic).
‖ Coal, iron, and steel. ¶ Textiles.

32. Although, generally speaking, the sample figures and *The*

Economist figures move in the same general direction so far as payments on ordinary shares are concerned—that is, payments are high in the same industries and low in the same industries—payments are, with a few exceptions, higher in the sample figures than in those given by *The Economist.* Of the thirteen industries in which comparison is possible, only in the case of one industry (stores) does *The Economist* figure exceed the sample figure, and this may be due to a different class content. In the sample, the 21 'stores' included 13 'drapers and general stores' and 8 'departmental stores', and the drapers in particular suffered more severely from the slump and made a slower recovery than the other companies in this group.

33. In the case of payments on 'prior' or preferred shares, the sample figures are higher than or equal to *The Economist* figures in 10 out of 13 cases. In payments on debentures, the sample figures are higher than or equal to *The Economist* figures in 8 out of 13 cases. (The figures for all industry do not bear comparison because of the large influence of the railways on the sample data.)

34. This may indicate that Crum's[1] conclusion for the United States, that the rate of profit increased with the size of the company (up to giant corporation size, where there was some indication of reversal of trend in the American data), might also be true for Great Britain, since the sample companies are, on the whole, larger than *The Economist* companies.

35. On the other hand, the difference may be due to the fact that the sample companies are all long-established companies whose actual capital in most cases is proportionately higher than that indicated by the nominal capital figures. This is no more than the slenderest of threads.

36. Perhaps the most surprising thing about the figures in Table III is the low rate of profit earned. The average true rate of profit (taking the median value, allowing for the bias in the sample due to the weight given to railways, and assuming that over- and under-capitalization were approximately equal)[2] was probably between 8 and 9 per cent. in the sample. If we take earnings gross of tax for ordinary shares alone, the true rate of profit was certainly rather higher, perhaps between 13 and 16 per cent.

37. Since the sample consists of large, well-established companies, the rate of profit in industry as a whole was probably somewhat lower than 8 to 9 per cent., although 1935 was a comparatively good year. It is suggested tentatively that in industry as a whole profits were as high as 10 per cent. in 1929 and as low as 5 per cent. in 1932.

---

[1] W. L. Crum, *Corporate Size and Earning Power*, Harvard, 1939, p. 23.

[2] Under-capitalization probably predominates, and the rate of profit would be correspondingly lower.

38. Variations between industries in 1935 ranged from the 23·3 per cent. earned in the tobacco industry, and the high returns in the motor, drink, chemical, building, and light engineering industries, to the low returns in the shipping (2·1 per cent. on total nominal capital and nothing for ordinary shares) and shipbuilding trades, railways, and textiles. These differences would undoubtedly be smaller were we able to estimate the actual capital investment. The high profits in the tobacco industry are probably explained in part by monopoly interests. They are certainly not altogether explained away by the extent to which nominal capital under-represented actual capital. But in the motor, light engineering, and brewing industries, and to a less extent in many other industries, nominal profit rates were higher than real profit rates, judging by the amounts set aside in undistributed profits. Low profits in shipping, shipbuilding, railways, and textiles are explained in part by over-capitalization.

39. Consumer goods industries including such durable consumer goods as cars and houses earned higher profits on the whole than capital goods industries, with the exception of the expanding motor, chemical, and light engineering industries.

40. Epstein[1] has suggested that there is little evidence in practice of the theoretically postulated tendency for profits to become equal in all industries, and cites as evidence the great variations in rates earned in different American industries. But as M. Tappan Holland in her review of the book pointed out,[2] the existence of very different profit rates does not invalidate the doctrine that under perfect competition profits will tend to equality. Excluding tobacco, real profits in British industry in 1935 varied perhaps between 3 and 15 per cent. Between 1935 and 1938 the labour force in several of the high-profit industries—motors, light engineering, amusement, chemicals—expanded considerably. In motors it increased by 39 per cent., in light engineering and amusement by 28 per cent., and in chemicals by 16 per cent. It is true that an even greater proportionate increase in employment took place in the shipbuilding industry in the same period, and there were increases of the same order on the railways and in the manufacture of locomotives and wagons. But recovery was late in these three industries, and for years after the First World War the railways stopped recruitment. In the same period the numbers employed in textiles fell by 11 per cent., and numbers in clothing and shipping showed little or no change, although the numbers employed generally had increased.

41. Restrictive practices are bound to distort the structure of profit rates, and competition was far from perfect. Nevertheless, the prosperous industries appeared to be attracting capital investment on the whole, and it is well known that the depressed industries certainly were not.

---

[1] R. C. Epstein, *Industrial Profits in the United States*, National Bureau of Economic Research, New York, 1934, p. 88 et seq.

[2] *Economic Journal*, 1936, p. 117 et seq.

7. Dividend Policies

42. The large part played in community saving by undistributed profits makes a study of directors' dividend policy—or, the complement of their dividend policy, saving policy—of the utmost importance.

43. Unfortunately the data used by the investigator yielded no information about what directors did with undistributed profits, though this is the most important consideration so far as trade-cycle theory is concerned. Undistributed profits can be kept as idle cash balances, temporarily invested in securities or stocks, invested more permanently in fixed assets, or used to increase working capital. If a business is expanding or trade is improving, larger amounts of working capital will be needed to finance the larger wages fund and stock-in-trade carried without any initiating decision by directors. If the policy, in relation to the use to which undistributed profits are put, pursued by directors changes over the period of the trade cycle—as it probably does—company savings can do much to accentuate the upward and downward phases. For instance, if a company pursues a policy of maintaining stable dividend payments, as many undoubtedly do, it may decide, once the boom has broken, to hold its reserves liquid lest it have to draw on reserves in the following year. If a great many companies do this, the consequent hoarding of cash and the cancellation of investment plans will accentuate the downward swing. Conversely, a policy of expanding the business, of investing in fixed assets, may have the opposite effect in boom years. On the other hand, a stable dividend policy does maintain consumers' purchasing power in the slump to some extent, since many companies are forced to draw on reserves.

44. Some light might be shed upon this problem by a study of changes from year to year in liquid assets as shown in the balance-sheet. All this investigation did was to indicate what dividend policies directors pursued and what proportion of earnings they saved in one form or another.

45. Company directors may have no dividend policy at all, but decide quite arbitrarily each year how much they distribute to ordinary shareholders and how much they save out of the residual profit which is left after all prior payments have been made. If, consciously or unconsciously, they do pursue some policy, it might take a number of different forms. They might save nothing at all and allow dividends to fluctuate directly with earnings, but in fact they do not do so. They might distribute less as earnings rise—a policy which has something to be said for it since it would mean that ordinary shareholders' income was more than maintained

during the slump—but there is no evidence in favour of such a theory. The three most likely policies to be pursued are as follows:

(*a*) A fixed amount put to reserve, and the rest distributed.

(*b*) Stable dividend payments, and the rest saved.

(*c*) Some policy between these two. That is to say, directors may have some conventional dividend payment in mind and, if anything above the amount necessary to cover this is earned, they may save a proportion of it and distribute the rest.

46. These three policies are numerically and figuratively illustrated in Table IV and Fig. 1 below. All amounts are net (i.e. after tax) because it is almost certainly net undistributed profits which directors will have in mind. They may compare these with gross payments on ordinary shares, but with the degree of fluctuation in income tax between 1924 and 1935 this will have little effect on the conclusions drawn.

TABLE IV

Probable Dividend Policies

(*a*)			(*b*)			(*c*)		
Stable undistributed profits			*Stable dividends*			*A conventional dividend plus (or minus) a proportion of earnings over (or under) the 'conventional' amount*		
Net earnings for ordinary shares	*Net undistributed profits*	*Net payments on ordinary shares*	*Net earnings for ordinary shares*	*Net undistributed profits*	*Net payments on ordinary shares*	*Net earnings for ordinary shares*	*Net undistributed profits*	*Net payments on ordinary shares*
£ mn.	£ mn.	£ mn.	£ mn.	£ mn.	£ mn.	£ mn.	£ mn.	£ mn.
5	2	3	5	1	4	5	1·5	3·5
4	2	2	4	0	4	4	1·0	3·0
10	2	8	10	6	4	10	4·0	6·0
7	2	5	7	3	4	7	2·5	4·5
7	2	5	7	3	4	7	2·5	4·5
8	2	6	8	4	4	8	3·0	5·0
8	2	6	8	4	4	8	3·0	5·0
6	2	4	6	2	4	6	2·0	4·0
3	2	1	3	−1	4	3	0·5	2·5
8	2	6	8	4	4	8	3·0	5·0

47. It will be seen that, if undistributed profits are plotted against total earnings for ordinary shares and a stable savings policy is pursued, the points will form a straight line parallel to the x-axis (Fig. 1, *a*). If stable dividends are the aim, the points will form a straight line at an angle of 45 degrees (if the scales on the x- and y-axes are the same) to the x-axis and cutting the x-axis at the point representing the amount

paid out in dividends (Fig. 1, *b*). If the policy is something between these two, the curve will be flatter and cut the *x*-axis nearer the origin. Thus in case (*c*) above, there was a 'conventional' dividend payment of £2 million. If anything above £2 million was earned, half of the excess was saved and half of it was paid out in dividends. If anything below £2 million had been earned, ordinary shareholders would not have suffered the full loss because reserves would have been drawn upon to the extent of 10*s*. in the £ for the amount by which earnings did not come up to conventional dividends.

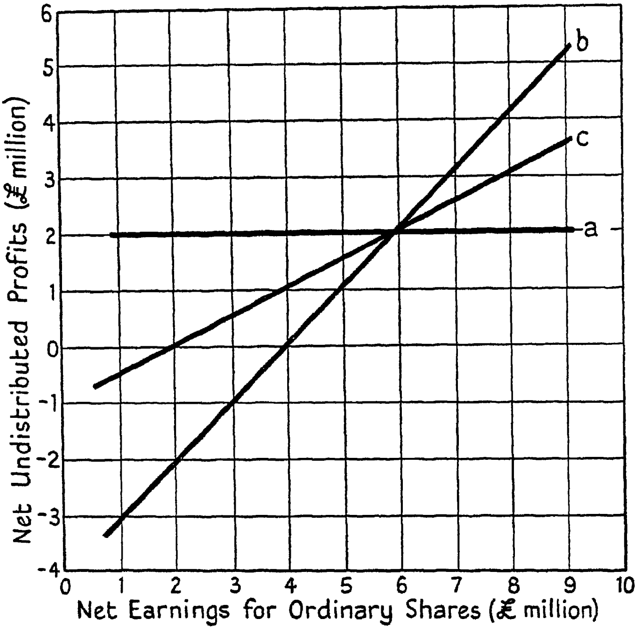

FIG. 1. Probable dividend policies

48. In case (*c*), of course, the assumption is made that the relationship is linear, i.e. that there is no increasing or decreasing marginal propensity to save as earnings increase. From the evidence collected this appears to be a reasonable assumption. The important point to notice here is that exactly the same result would obtain if cases (*a*) and (*b*) were added together and divided by two. In practice, case (*c*) appears to be the most common form which dividend policies take, but we cannot exclude the possibility that some companies in each sample may be pursuing policy (*a*). The *results* of dividend policy differ from industry to industry, but it is doubtful whether there will be any particular *practice* common to all firms in a given industry and distinct from the practice common to all firms in another industry.

49. In almost every industrial sample examined one or two straight lines fitted the points for each of the years from 1924 to 1935 fairly well. Inductively we may say that the relationship between earnings for dividend and undistributed profits appears to be linear in form.

50. As a general rule it may be suggested that given earnings for ordinary dividends of y, directors paid out in ordinary dividends a fixed amount a (here called the 'conventional' dividend payment) plus a fixed proportion b of the residual amount $(y-a)$, whatever the sign of $(y-a)$. Thus, if x were the amount of ordinary dividend paid,

$$x = a+b(y-a).$$

51. This information is summarized in Table V, which gives the value of the parameters a and b in the above equation. The straight lines were fitted by eye and are not mathematically precise, but the nature of the data does not warrant more exact formulation.

52. The 'conventional' dividend payment, a, is represented as a rough percentage of the paid-up ordinary capital in 1935 and is no more than very approximate. Nevertheless it is noteworthy that in many industries, in which vastly different dividend policies were pursued, 5 per cent. seems to have been the 'normal' figure which directors had in mind, at least before the slump. This is true of miscellaneous and electrical engineering, locomotives and wagons, and textile manufacturers and merchants. Below 5 per cent. we find shipbuilding, iron and steel, coal, large engineering, chemicals (1924–8), docks and wharves, shipping, railways, clothing, and hotels and restaurants. The lowest values of a before 1930 are found among large engineering firms, in railways, shipbuilding, docks and wharves, hotels and restaurants—all around 3 per cent. After 1930 the value of a fell as low as 1 per cent. in shipping and railways. In the motor industry, between 6 and 7 per cent. seems to have been 'normal'. The highest values, from 7 to 9 per cent., were found in stores, drink, leather, and food.

53. The value of b varied between less than 1*s*. in the £ and 10*s*. in the £, or between 0·00 and 0·50, with values even higher in one or two exceptional cases.

54. The lower the value of b, the more stable were dividend payments, and many directors went to great lengths by way of drawing on reserves in bad years in order to keep dividend payments relatively stable.

55. Industries in which stable dividend policies were pursued (a value for b below 0·25) included coal from 1927 to 1932, iron and steel from 1924 to 1929, miscellaneous engineering (in spite of the fact that positive earnings for ordinary shares were made in only 4 years out of the 12), shipbuilding (again although positive earnings were made in only 4 years),

TABLE V

Dividend Policies in Practice

Industry	*Period*	*'Conventional' dividend payment*[1] *a*	*Slope of curve*[1] *b*	*Remarks*
1	2	3	4	5
All industry .	1924–30	6·5	0·33	. .
	1931–5	4·5	0·47	. .
Food . .	1924–7	7·0	0·10	A close fit in the latter period.
	1928–35	9·0	0·00	. .
Drink . .	1924–35	8·5	0·50	A tendency towards stable dividends in first 2 or 3 years, otherwise a close fit.
Tobacco . .	1924–35	. .	. .	A roughly constant amount saved; see text.
Paper . .	1926–35	4·5	0·62	Wide scatter and considerable fluctuations in payments; see text.
Stores . .	1924–35	9·0	0·40	Considerable scatter.
Hotels and Restaurants	1924–35	3·0	0·55	A tendency towards stable dividends in first few years.
Amusement .	1924–35	4·5	0·25	. .
Leather . .	1924–35	8·8	0·25	Except for 1930 and 1931, a good fit.
Clothing . .	1924–35	4·0	0·45	Considerable scatter and 1924, the best year, proportionately more saved.
Building . .	1926–34	5·8	0·40	1924 and 1925 are well above and 1935 well below the line; otherwise a good fit.
Non-commercial Motors	1924–33	7·0	0·00	Last 2 or 3 years indefinite, but payments rose.
Commercial Motors	1924–7	6·5	0·05	. .
	1928–35	9·0	0·15	. .
Vehicle Accessories . .	1924–35	5·8	0·20	Considerable scatter.
Electrical Engineering	1924–32	5·0	0·00	Considerable scatter.
	1933–5	5·0	0·42	. .
Chemicals . .	1924–8	4·0	0·50	Nominal capital may have increased between the first and second periods. Decided to save loss in the latter period.
	1929–32	6·0	0·50	
	1933–5	6·0	0·73	
General Engineering	1924–31	7·5	0·30	. .
	1932–5	4·0	0·30	. .
Miscellaneous Engineering	1924–8	5·0	0·05	Mostly a process of drawing on reserves to maintain stable dividends.
	1929–35	2·8	0·20	
Large Engineering	1924–35	2·5	0·30	. .
Iron and Steel .	1924–9	4·2	0·05	Stable dividends in first period.
	1930–5	3·5	0·42	. .
Coal . . .	1924–6	4·5	0·50	Considerable scatter and difficult to interpret.
	1927–32	1·5	0·10	
	1929–35	1·5	0·50	

---

[1] $x = a+b(y-a)$, where x = net payments on ordinary shares,
y = net earnings for ordinary shares,
a = a fixed amount, represented in Table V as a rough percentage of paid-up ordinary capital in 1935,
and b = a fixed percentage (e.g. 0·40 = 40 per cent. or 8*s*. in the £) of $y-a$.

Table V (cont.)

Industry	*Period*	*'Conventional' dividend payment*[1] *a*	*Slope of curve*[1] *b*	*Remarks*
1	2	3	4	5
Railways	1924–6	4·0	0·25	In 9 of the 12 years reserves were drawn upon, though payments were progressively reduced.
	1927–30	2·5	0·55	
	1931–5	1·0	0·55	
Locomotives and Wagons	1924–35	5·0	0·55	A good fit, though some tendency for *a* and *b* to fall in the last 3 years.
Shipping	1924–32	4·5	0·40	Reserves were drawn upon in all but 1 of the last 6 years, and positive earnings for ordinary shares were made only in 1 (1934) of the last 5 years.
	1933–5	1·0	0·00	
Docks and Wharves	1924–35	3·0	0·55	A good fit.
Shipbuilding	1924–35	3·0	0·15	A good fit in spite of the fact that no net earnings for ordinary dividends were made in 8 of the 12 years and amounts were put to reserve only in 1929 and 1935.
Textile Manufacturers	1924–35	5·0	0·35	Considerable scatter. From 1924 to 1928 the marginal propensity to save was higher when earnings were high. From 1930 to 1933 the marginal propensity to save decreased as earnings increased.
Textile Merchants	1924–35	5·0	0·40	Very similar to textile manufacturers except that from 1931 to 1935 the marginal propensity to save increased with earnings.

[1] See opposite.

electrical engineering from 1924 to 1933, all the motor trades, building from 1924 to 1931, and food.

56. Considerable fluctuations in payments (a value for *b* above 0·5) were made in locomotives and wagons, chemicals from 1933 to 1935, docks and wharves, railways from 1927 to 1935, and hotels and restaurants.

57. A change of policy appears to have taken place in many industries during the slump. This appears to be markedly true of the sample as a whole. In general, the change took the direction of a lower 'conventional' dividend and a flattening out of the savings curve, indicating a greater willingness to allow dividend payments to fluctuate more with earnings. In other words, the value of *a* fell while the value of *b* rose. Coal, iron and steel, miscellaneous engineering, commercial motors, railways, and food are among the industries in which two separate curves, one for the first half of the period and another for the second half, fit the facts well. In electrical engineering and chemicals the conventional rate remained the same but the amount distributed fluctuated more, while in

general engineering the rate was reduced while the slope remained the same.

58. The tobacco and the paper industries are the only ones in which no curve can conveniently be fitted. In the tobacco industry it would appear that a roughly constant amount was set aside to reserve every year and the rest distributed, with the notable exceptions of 1926, when more than usual was set aside, and 1933, when reserves were drawn upon. Generally speaking the amount put to reserve was lower after the slump than before it.

59. In the paper industry a smooth curve, indicating an increasing marginal propensity to save as earnings increased, fits the points for the years 1927 to 1931, but for the remaining years no particular policy is apparent, though it should be remembered that this industry covers companies producing a great variety of products, from cartons to comic papers.

60. In one or two other cases a curve of a higher power seems to fit the data better than a straight line, though the curve is not always of the shape that might be expected. Among textile manufacturers, for example, marginal propensity to save increased with earnings from 1924 to 1928, but decreased from 1930 to 1933. Among textile merchants, marginal propensity to save increased with earnings from 1924 to 1927 and from 1932 to 1935, but at a faster rate in the latter period. In building, which is a greatly diversified industry in the sample and hence one in which a wide scatter might be expected, a smooth curve indicating increasing marginal propensity to save can be drawn through the points over the period 1925–34, excluding 1926. There is also some indication of an increasing marginal propensity to save in clothing and hotels and restaurants.

61. We can easily fall into the error of reading more into the data than is to be found there, but, although in most cases there is some degree of scatter, the linear relationship between earnings for residual shares and undistributed profits is surprisingly common. When it is borne in mind that many transient factors will weigh with directors when they are deciding dividend policy in any one year, the consistency of the data becomes the more surprising.

62. It should be remembered that these samples are made up of companies some of which suffered losses in many of the years covered, and it is theoretically possible that the policies of directors might be different according to whether earnings for ordinary dividend were positive or negative. It is, however, abundantly clear that in bad years most directors will draw very heavily on reserves, even if there are several bad years running, in order to maintain some dividend payments, and that these dividend payments appear to have some definite relationship to those

made in good years. The presumption is that no striking difference of policy exists between firms suffering losses and those making profits.

8. Savings Achieved

63. Table VI presents the data concerning undistributed profits in a different form. Column 2 expresses net undistributed profits over the period 1924–35 as a percentage of the total net earnings for ordinary shares in this period. In columns 3 and 4 the percentage is given for the periods 1924–9 and 1930–5 respectively, and in column 5 total net earnings for ordinary shares from 1930 to 1935 are expressed as a percentage of the total net earnings for the first 6 years.

TABLE VI

Percentage of Net Earnings for Ordinary Shares Saved

Industry	*Net undistributed profits as percentage of total net earnings for ordinary shares*			*Net earnings for ordinary shares 1930–5 as percentage of 1924–9*
	1924–35	*1924–9*	*1930–5*	
(1)	(2)	(3)	(4)	(5)
All industry . . .	10·5	13	7	70
Non-commercial Motors .	61·6	66	57	113
Commercial Motors .	45·4	65	−29	27
Amusement . . .	37·5	38	36	72
Motor Accessories . .	36·7	45	24	59
Large Engineering . .	34·5	43	24	77
Clothing . . .	31·6	40	1	28
Electrical Engineering .	30·4	41	19	87
Hotels and Restaurants .	30·0	32	26	69
Paper	28·5	32	25	82
Brewing . . .	25·5	26	24	82
Food	23·6	27	20	105
Building . . .	23·0	32	13	88
Docks and Wharves .	21·4	24	17	57
Stores	19·8	26	9	60
Chemicals . . .	17·6	25	11	102
Textile Manufacturers .	14·3	24	−22	27
Leather . . .	14·3	27	−8	54
Iron and Steel . .	11·2	−20	36	125
General Engineering .	9·2	18	−7	58
Tobacco . . .	7·3	11	4	96
Shipping . . .	5·5	26	−1,063	2
Locomotives and Wagons	−1·9	10	−41	31
Coal	−8·0	−90	27	246
Textile Merchants . .	−57·4	−10	No net earnings for dividend	. .
Railways . . .	−83·2	−89	−79	31
Shipbuilding . . .	No net earnings for ordinary shares; reserves drawn on to 2·5 times extent of losses.			
Miscellaneous Engineering	No net earnings for ordinary shares; reserves drawn on to 4·5 times extent of losses.			

64. It will be noticed that, with the exception of a few industries, earnings in the second 6 years were lower, often considerably lower, than they were in the first. As has been suggested above,[1] the coal and iron and steel figures are no doubt largely explained by the disastrous effects upon profits of the coal lock-out in the earlier period. Non-commercial motors, chemicals, and food are the only other exceptions, and these were expanding and profitable industries. The fall in shipping, commercial motor, and textile earnings is noteworthy, while the sample of shipbuilding and miscellaneous engineering firms made no net earnings over the whole period.

65. The striking thing about these figures is the immense differences between industries in the savings actually achieved, quite apart from the policy pursued. Over the period as a whole a high proportion of earnings was saved in the motor industry, over 60 per cent. in non-commercial motors and 45 per cent. in commercial motors. Over 30 per cent. of earnings were saved in the amusement, vehicle accessories, large engineering, clothing, electrical engineering, and hotel and restaurant industries, and over 20 per cent. in paper, brewing, food, building, and docks and wharves. On the other hand, the railways, textile merchants, and, to a lesser extent, coal producers and locomotive and wagon manufacturers actually disbursed more than they earned. In shipbuilding and among the small sample of miscellaneous engineers no earnings for ordinary shares were made at all on balance, though reserves were drawn upon considerably in excess of the losses suffered.

66. Equally striking are the differences between the first and second halves of the period. Before 1930 most industries saved at least a quarter of earnings for ordinary shares. After 1930 this was true of not more than half a dozen—non-commercial motors, amusement, hotels and restaurants, paper, iron and steel, and coal. In every case, with the exceptions of coal and iron and steel, the savings ratio was lower in the second half of the period than in the first half, even where earnings were higher. Disbursements were greater than earnings among commercial motor, textile manufacturing, leather, general engineering, shipping, locomotive and wagon, and railway companies, as well as among textile merchants, in shipbuilding and miscellaneous engineering, where no net earnings for ordinary dividend were made.

9. Shareholders' Incomes

67. There is no space in this article to detail some of the interesting facts relating to particular industries which came to light in the course of this investigation. Only the main findings of general application are here presented. In conclusion we may note that, as we should expect and in spite of the cushion afforded by undistributed profits and reserves, the burden of fluctuations in income in the community was undoubtedly borne by the risk-taker, i.e. by the ordinary shareholder. His income fluctuated between two and three times as much as the national income, between four and five times as much as wages, and over five times as much as salaries. Generally speaking, profits and shareholders' incomes increased faster than other incomes in times of improving trade and fell faster than other incomes in times of depression.

---

[1] p. 166, para. 25.

APPENDIX

A Note on Method

68. The Companies Acts previous to that of 1947 did not specify what any statement of profit and loss should contain. Even from a carefully chosen sample of the kind used here it is not possible to obtain a completely accurate picture. Published figures of past profits may differ from those agreed with the Inland Revenue in a number of ways.

69. First of all, they exclude the tax due upon the year's profits, with the exception of the tax on dividends where dividends are shown gross, as they often are.

70. Second, the directors may have charged more or less for the purpose of depreciation than is allowed by the Inland Revenue authorities. In good years at least, since the latter do not allow for some types of obsolescence which business men agree in reckoning as costs, the shown profit will probably be less than the Inland Revenue estimate. In bad years fixed rates of depreciation will continue to be applied, but these may not be on the same basis as those allowed by the Inland Revenue.

71. Third, the tax provision made before showing the results of the year's activities did not necessarily bear any relation to the earnings of the period covered by the accounts. Directors might have charged against each year the tax actually due in that year, and if this were done the method adopted by the investigator for estimating income tax would overestimate profits when profits were rising and vice versa. Alternatively they might have used the income-tax provision as an account for making allocations to, or drawings from, reserve. If this were done, there is no way of allowing for it, any more than there is with operations of the same kind on the depreciation account. It has been assumed that they did neither of these things, but made provision for income tax on the basis of the year's earnings. If the other policies were adopted by some directors, we may take some consolation in the thought that the effects of the one would presumably be in a direction opposite to those of the other.

72. Fourth, many companies, in the period under review, made provision for reserves before showing profits. This might be done by writing down assets or investments, by using the tax reserve, or by including hidden reserves under such omnibus headings as 'Sundry Creditors and Credit Balances' and 'Provision for Contingencies'. (Generally speaking, these practices were thought to be on a smaller scale among industrial firms than among banks and insurance companies.) Conversely, items from hidden reserves might be included in the shown profit. These practices again tend to reduce the profit shown in good years and increase it in bad years, and to obscure the earning power of the company in the long period.

73. Fifth and last, all of the interest paid might not be shown. Debenture interest was fairly well shown and, where not shown but obviously paid, could usually be estimated. Mortgage interest was, however, rarely shown by industrial companies, being tied up quite often with the tax reserve or some other item. Since the balance-sheets did not state the percentage rate of mortgage interest, any gaps here could not be filled in.

74. The second, fourth, and fifth factors suggest that profits shown to shareholders in the past were less than actual profits, particularly in good years. On the other hand, the 510 companies in the sample examined, since they were companies for which information was regularly available, probably earned profits which were higher than those of most other companies. Unfortunately there is no way of telling how far these different tendencies cancel out.

75. In correcting for income tax, no allowance can be made for tax already deducted from receipts on securities held because these were not shown separately. Some adjustments could be, and were, made to the shown 'earned for dividends' figure. Allocations to mortgage or debenture redemption and amortization accounts were added back where they were shown as a prior charge, since they were considered allocations to reserve. Similarly, sums written off investments, initial outlay, capital and development expenses, &c., were added back where shown as a prior charge, unless it was clear that the writing off had been made necessary by a shrinkage in the current value of the assets. Where no evidence to the contrary existed, amounts written off were taken as allocations to reserve.

76. The shown figure, amended in this way, was then assumed to be the correct one. To obtain, from the shown figures, total *net earnings for dividend* for any group of companies, the following totals were calculated:

(*a*) The total 'earnings for dividend' shown.

(*b*) The total 'net preferred dividend payments' and the total 'net ordinary dividend payments' for those companies which showed such payments free of tax.

(*c*) The total 'gross preferred dividend payments' and the total 'gross ordinary dividend payments' for those companies which showed such payments gross.

The tax paid on the totals paid gross was then calculated at the standard rate applicable and deducted from the total of shown 'earnings for dividend', yielding the best approximation to total net earnings for dividend which can be obtained.

77. To obtain from this figure total *gross earnings for dividend*, an estimate has to be made of income tax actually paid on net earnings. For this purpose, we must *add* back to total net earnings the sum of losses shown in the published accounts for the year in question, and *add* further the tax deducted from any dividends shown gross where those dividends have not been earned.

78. The point here is that if a company has not made a profit, it has no tax to pay. If it shows dividends gross, then the part of tax not recovered by deduction from the coupons will be included among trading expenses before arriving at trading profit. If, however, dividends are paid wholly out of reserve, 'the part of tax not recovered by deduction from the coupons' will be negative, i.e. the trading account will be *credited* with the amount deducted from coupons. The result which the trading account apparently yields will therefore be improved by the amount of the tax. Hence a company has made a loss, even if its trading account shows a profit, if that profit is less than the amount deducted from coupons and credited to the trading account. The criterion is whether the shown earnings for dividend are less than the tax on dividends shown gross. The amount by which losses are greater than they appear at first sight on this account, and hence the sum which is added to total net earnings in addition to apparent losses can be calculated only by detailed inspection of the accounts of individual companies.

79. From the sum thus obtained, we must *deduct* the loss allowance due to any firm for losses made in previous years. For example, if in the previous year a firm made a loss of £400 and in the year in question showed earnings for dividend amounting to £3,200, earnings for dividend on which tax was paid would be £2,800 and £400 must be deducted from the above total. If, however, the loss in the previous year was £3,600, the earnings for dividend shown—£3,200—would be deducted, since they would be free of tax, and the balance of £400 would be carried forward to be deducted from earnings for dividend in the next year if these earnings were positive.

80. Lastly, we must *deduct* any other items included in earnings for dividend which are free of tax—mainly tax refunds. These amounts also have to be determined by careful inspection of individual accounts.

81. The total thus obtained will be the total earnings for dividend on which tax was paid, so far as these can be determined, *after the deduction of tax.* From this we can calculate the amount paid in tax, e.g. if income-tax is 5*s*. in the £, by taking one-third of this total. By adding tax paid to total net earnings for dividend, we obtain total gross earnings for dividend. By adding gross interest payments to gross earnings for dividend, we obtain gross profits. By deducting total net dividends from net earnings for dividend, we obtain net undistributed profits. These are the various figures which have been used in this paper, in addition to the capital figures which were estimated for 1935–6 only.

82. One further difficulty needs to be noted. Company accounts may be presented at any time during the year. The year taken runs from 1 July to 30 June and, since profits are earned in a previous period, the figures are taken as applicable to the calendar year beginning and ending six months previously. But the balance-sheet and profit and loss accounts may be made up in any month in the year. Consequently, since there is sometimes a favoured quarter in which accounts are presented in particular industries, there may be some irregularity in the way in which different industries contribute to the various periods. The examination of quarterly totals in a number of industries does not suggest that this possibility affects the results so seriously in practice as it might in theory.

Acknowledgements

This article is based on a small part of the material collected by Mr. P. W. S. Andrews in 1938–9 from the analyses of balance-sheets and profit and loss accounts published by Moody's Services, Ltd. I cannot adequately express my gratitude to Mr. Andrews and Professor E. H. Phelps Brown, both of whom gave me very practical assistance. My thanks are also due to the Oxford Economists' Research Group, under whose auspices Mr. Andrews worked, for giving me full permission to make whatever use of the material I liked, and to Moody's Services, Ltd., for permission to make the information available. Such views as are expressed are, of course, my responsibility alone.

SELWYN HOUSE, ENDSLEIGH STREET,
LONDON, W.C. 1.

companies, being tied up quite often with the tax reserve or some other item. Since the balance-sheets did not state the percentage rate of mortgage interest, any gaps here could not be filled in.

74. The second, fourth, and fifth factors suggest that profits shown to shareholders in the past were less than actual profits, particularly in good years. On the other hand, the 510 companies in the sample examined, since they were companies for which information was regularly available, probably earned profits which were higher than those of most other companies. Unfortunately there is no way of telling how far these different tendencies cancel out.

75. In correcting for income tax, no allowance can be made for tax already deducted from receipts on securities held because these were not shown separately. Some adjustments could be, and were, made to the shown 'earned for dividends' figure. Allocations to mortgage or debenture redemption and amortization accounts were added back where they were shown as a prior charge, since they were considered allocations to reserve. Similarly, sums written off investments, initial outlay, capital and development expenses, &c., were added back where shown as a prior charge, unless it was clear that the writing off had been made necessary by a shrinkage in the current value of the assets. Where no evidence to the contrary existed, amounts written off were taken as allocations to reserve.

76. The shown figure, amended in this way, was then assumed to be the correct one. To obtain, from the shown figures, total *net earnings for dividend* for any group of companies, the following totals were calculated:

(*a*) The total 'earnings for dividend' shown.

(*b*) The total 'net preferred dividend payments' and the total 'net ordinary dividend payments' for those companies which showed such payments free of tax.

(*c*) The total 'gross preferred dividend payments' and the total 'gross ordinary dividend payments' for those companies which showed such payments gross.

The tax paid on the totals paid gross was then calculated at the standard rate applicable and deducted from the total of shown 'earnings for dividend', yielding the best approximation to total net earnings for dividend which can be obtained.

77. To obtain from this figure total *gross earnings for dividend,* an estimate has to be made of income tax actually paid on net earnings. For this purpose, we must *add* back to total net earnings the sum of losses shown in the published accounts for the year in question, and *add* further the tax deducted from any dividends shown gross where those dividends have not been earned.

78. The point here is that if a company has not made a profit, it has no tax to pay. If it shows dividends gross, then the part of tax not recovered by deduction from the coupons will be included among trading expenses before arriving at trading profit. If, however, dividends are paid wholly out of reserve, 'the part of tax not recovered by deduction from the coupons' will be negative, i.e. the trading account will be *credited* with the amount deducted from coupons. The result which the trading account apparently yields will therefore be improved by the amount of the tax. Hence a company has made a loss, even if its trading account shows a profit, if that profit is less than the amount deducted from coupons and credited to the trading account. The criterion is whether the shown earnings for dividend are less than the tax on dividends shown gross. The amount by which losses are greater than they appear at first sight on this account, and hence the sum which is added to total net earnings in addition to apparent losses can be calculated only by detailed inspection of the accounts of individual companies.

79. From the sum thus obtained, we must *deduct* the loss allowance due to any firm for losses made in previous years. For example, if in the previous year a firm made a loss of £400 and in the year in question showed earnings for dividend amounting to £3,200, earnings for dividend on which tax was paid would be £2,800 and £400 must be deducted from the above total. If, however, the loss in the previous year was £3,600, the earnings for dividend shown—£3,200—would be deducted, since they would be free of tax, and the balance of £400 would be carried forward to be deducted from earnings for dividend in the next year if these earnings were positive.

80. Lastly, we must *deduct* any other items included in earnings for dividend which are free of tax—mainly tax refunds. These amounts also have to be determined by careful inspection of individual accounts.

81. The total thus obtained will be the total earnings for dividend on which tax was paid, so far as these can be determined, *after the deduction of tax*. From this we can calculate the amount paid in tax, e.g. if income-tax is 5*s*. in the £, by taking one-third of this total. By adding tax paid to total net earnings for dividend, we obtain total gross earnings for dividend. By adding gross interest payments to gross earnings for dividend, we obtain gross profits. By deducting total net dividends from net earnings for dividend, we obtain net undistributed profits. These are the various figures which have been used in this paper, in addition to the capital figures which were estimated for 1935–6 only.

82. One further difficulty needs to be noted. Company accounts may be presented at any time during the year. The year taken runs from 1 July to 30 June and, since profits are earned in a previous period, the figures are taken as applicable to the calendar year beginning and ending six months previously. But the balance-sheet and profit and loss accounts may be made up in any month in the year. Consequently, since there is sometimes a favoured quarter in which accounts are presented in particular industries, there may be some irregularity in the way in which different industries contribute to the various periods. The examination of quarterly totals in a number of industries does not suggest that this possibility affects the results so seriously in practice as it might in theory.

Acknowledgements

This article is based on a small part of the material collected by Mr. P. W. S. Andrews in 1938–9 from the analyses of balance-sheets and profit and loss accounts published by Moody's Services, Ltd. I cannot adequately express my gratitude to Mr. Andrews and Professor E. H. Phelps Brown, both of whom gave me very practical assistance. My thanks are also due to the Oxford Economists' Research Group, under whose auspices Mr. Andrews worked, for giving me full permission to make whatever use of the material I liked, and to Moody's Services, Ltd., for permission to make the information available. Such views as are expressed are, of course, my responsibility alone.

SELWYN HOUSE, ENDSLEIGH STREET,
LONDON, W.C. 1.

SOME THOUGHTS ON THE NATURE OF INTEREST

By E. VICTOR MORGAN

THE division among English economists between those who regarded the rate of interest as determined by the supply and demand for money and those who thought of it as depending on the supply and demand for savings goes back at least to the end of the seventeenth century. The latter view secured the authority first of David Hume and then of Smith and Ricardo, and became generally accepted among academic economists. The former, however, persisted especially among practical bankers with experience of the consequences of gold movements and variations in Bank of England policy. The growing concern with fluctuations in income and employment produced renewed speculation about interest and culminated in Keynes's *General Theory of Employment, Interest, and Money.* Keynes criticized the 'classical' view on the grounds that, unless the level of income were postulated, it left the volume of saving and the rate of interest indeterminate. Instead, we have in 'liquidity preference' the old monetary view of interest presented in a far more logical form than ever before and made, for the first time, respectable.

This theory depends on a sharp distinction between two decisions of the individual:

(*a*) to save or to spend, and
(*b*) to hold one's wealth in one form rather than another.

It also depends upon the existence of uncertainty, for if there were no uncertainty about future rates of interest there would be no advantage in holding money except in the amounts required for current transactions. The existence of uncertainty as to future rates of interest gives a motive for holding money as a precaution against depreciation in the capital value of debts due to a future rise in the rate of interest (Keynes's precautionary motive). Further, when there are differences of opinion as to what the future level of interest rates may be, there is a further motive for holding money by the bears (those who expect a rise in interest rates) in order to buy debts on more favourable terms if their expectations are eventually fulfilled (Keynes's speculative motive).

The theory has been criticized by Professor Hicks and others on the ground that it makes the very existence of the rate of interest depend upon uncertainty as to future rates of interest and therefore leaves interest 'hanging by its own bootstraps'. It seems to me that the underlying truth of this criticism is that the theory over-emphasizes the clarity of division between the decision to save or spend and that as to the form in which to hold one's wealth, and that, in considering the latter decision, it concentrates too exclusively on one choice—the choice between money and debts. A fully satisfactory theory should take account of the whole range of possible choices. There is no doubt that Keynes realized this, and chapter xvii of the *General Theory* is apparently an attempt to remedy the defect. Unfortunately the chapter is a very difficult one, and its position in the book suggests that it is a refinement rather than a fundamental part of the argument.

In order to analyse further the nature of the rate of interest we shall, therefore, reduce the money-or-debts choice to its true perspective by considering the whole range of alternatives open to the property holder. We shall also put away, for the present, the distinction between these decisions and the save-or-spend decision—useful as it is in other contexts. In reality this distinction is less sharp than it might appear, since the property holder can always 'dis-save', i.e. convert his property into cash and 'blue' it on current consumption. This fact, which was rather discounted before the war, is all too obvious to-day.

The holder of claims to wealth—whether received as current income or carried forward as property from the past—has before him a series of alternatives which can be divided into the following broad categories:

(*a*) Immediate consumption, i.e. the purchase of services and of goods which are used up within a short time in the course of a single act of self-satisfaction.

(*b*) The purchase of durable consumers' goods, i.e. goods which will continue to yield satisfaction over a long period.

(*c*) The purchase of debts of one form or another.

(*d*) The acquisition of stocks of goods intended not for personal use but for resale.

(*e*) The acquisition of fixed capital assets. The distinction between durable consumers' goods and capital goods is that the services performed by the latter are intended for personal use, while those rendered by the former are intended for resale.

(*f*) The holding of money.

Each of these categories, except the last, can be subdivided into an almost infinite number of subsections. The dividing-line between different categories is not always clear; for example, a bank deposit on current account could be regarded as money or as a debt payable on demand and backed by almost perfect security. This difficulty, however, will not affect our argument.

A person will be willing to acquire goods, debts, or money because of certain advantages which each brings him. These advantages will be:

(*a*) For services and non-durable consumers' goods, the immediate satisfaction derived from their use.

(*b*) For durable consumers' goods, the services which these goods may be expected to render during their life, discounted at a rate representing the purchasers' rate of time preference. For example, suppose that a motor-car is expected to yield £100 worth of motoring a year for five years, but its purchaser regards £100 in one year as equivalent to £105 in the next year. Then his demand price for the car will be:

$$£100+£100\times\frac{100}{105}+£100\times\frac{(100)^2}{105}+£100\times\frac{(100)^3}{105}+£100\times\frac{(100)^4}{105}$$

(*c*) For debts, the dividend received, less the value attached to the risk of loss and the inconvenience attending the transaction.

(*d*) For stocks of commodities, the convenience in working which they afford, plus any increase in value which may be anticipated, less carrying costs and risk pressure.

(*e*) For fixed capital, the prospective earnings of the instrument concerned. These will, in the short run, be quasi-rents determined by conditions of supply and demand for the final product and of other factors and without reference to the cost of production of the capital good. The advantages derived from fixed capital and stocks like the future services derived from durable consumers' goods will be discounted at a rate representing the holder's preference for present rather than future benefits.

(*f*) For money, although in the holding of a certain stock of money there will be great convenience in the making of current transactions, beyond a certain point the advantages from this point of view of further additions to the money stock may be expected to decline rapidly. Further advantages may, however, be provided by Keynes's liquidity motives.

Clearly the value set upon these advantages—at the margin—by any individual will depend upon the amount of the different types of good which he holds, and the schedule of marginal net advantages for any one good will differ from one person to another according to varying tastes and circumstances. We shall call the ratio marginal net advantage of A/marginal net advantage of B the marginal rate of substitution of A for B. So long as this ratio is greater than Price of A/Price of B, a purchaser will have an incentive to buy more of A and less of B and vice versa. Equilibrium can, of course, only be attained when the two ratios are equal.

The individual citizen has an unfettered choice between these alternatives (in the absence of legal controls). If he has money he can either keep it or spend it on any type of good or debt he chooses. If he has one type of good or debt he can sell it and buy another without (by his own act alone) causing any significant change in market prices. But clearly there is no such freedom for society as a whole. At any moment of time the total quantity of each type of good, of each type of debt, and of money in existence is fixed, and in most cases changes in that quantity take place only gradually; important exceptions to the second part of this statement are stocks of consumers' goods and materials, government bonds in wartime, and token money during an inflation.

We shall consider the operation of this freedom of individual choice in relation to three types of equilibrium:

A. Market equilibrium, where we take the stock of all goods as given.

B. Static equilibrium over a period of time with constant rates of output of each good and, implicit in this assumption, a stable population and zero net investment, and

C. Dynamic equilibrium with a changing rate of output and a changing rate of net investment. Here we postulate only that capital shall be used to optimum intensity, i.e. to the intensity at which the marginal net products of all other factors are equal to their prices, and that these prices shall be such as to leave none of the other factors involuntarily idle.

A. In considering the determination of prices in the condition of Marshall's 'market equilibrium', we assume first a barter economy in which we have consumer goods, both durable and non-durable, stocks of such goods and of materials, but neither money nor debts. Since the amounts of all types of good in existence are fixed, there can be no possibility of saving or dis-saving by the community as a whole. Any individual, however, can increase his holding of one asset and diminish his holding of another by doing a simple swap. In order that equilibrium may exist it is necessary that every individual should be in a position where his marginal rate of substitution is equal to the rate of exchange, for any pair of goods which he holds, and greater than the rate of exchange as between a good which he holds and one which he does not hold.

The way in which this works out in the non-durable consumers' goods market is simple. In dealing with durable consumers' goods, stocks, and fixed capital, the matter is complicated by the question of time preference.

An individual will presumably discount the expected benefits from different types of durable good at the same rate. This rate, however, will not be constant but will vary according to the proportion of the person's total resources which is locked up in durable goods. With a given amount of total resources we may expect that, as the amount of durable goods held increases, so the rate of discount required to compensate for the holding of a marginal unit will also increase. We shall thus have a schedule of rates of time preference applicable to varying amounts of durable goods. Moreover, it is not necessary that the rate of time preference will be a uniform amount per year. Indeed, in view of the uncertainties of human life, it is quite likely that benefits accruing in the remote future will be discounted at a higher rate per annum than benefits nearer to hand. If this is so, we shall have to think of a series of schedules corresponding to different degrees of futurity. For the sake of simplifying the following argument, however, we shall assume that benefits of differing degrees of futurity are, in fact, discounted at a uniform rate per annum. Finally, our time-preference schedules will, of course, differ from one person to another, according to the magnitude of their resources and their subjective attitude towards the future.

The marginal rate of substitution for any holder between any two lasting assets will be the ratios of their prospective net advantages discounted at the appropriate rate of time preference, and swaps will continue until these marginal rates of substitution are equal to exchange rates in all cases. But it is necessary also that equilibrium should exist between immediate consumption and lasting goods, and for this purpose it is necessary that the rate of discount of future benefits, which is implicit in the marginal rate of substitution between a durable and a non-durable good, should be such as to compensate the most reluctant holder of durable goods who is necessary in order to absorb the existing supply. We will call this the marginal rate of time preference. We thus reach the proposition that the 'price' of capital assets must equal the sum of their prospective net advantages discounted at the marginal rate of time preference.

Our marginal rate of time preference will depend on the value set by our community on future as compared with present benefits, upon the size and distribution of total resources, and upon the relative amounts of present and future benefits available. We may expect it to be positive, but it is quite possible to conceive of its being negative. We thus have something very like the rate of interest whose existence depends neither on saving nor on money, but simply on the existence of some goods yielding immediate benefits and some yielding deferred benefits, and on the two being regarded as not exactly equivalent.

It remains, while still considering market equilibrium, to introduce debts and money into the picture. Here it is necessary to make a further qualification about uncertainty. To begin with, we shall assume that there is no uncertainty about the future of the marginal rate of time preference.

The holder of a debt payable in kind expects to get the same advantage as if he held a piece of capital capable of producing the good concerned in similar quantities after allowing for all relevant costs. He will save himself the trouble and risk of direct ownership and will incur instead the risk of default. The prospective net advantages will both be discounted at the marginal rate of time preference, and the exchange rates and therefore the yield to new purchasers will represent the value set upon the difference between these two risks.

We now introduce money and have the relief of being able to talk about prices instead of rates of exchange. In the absence of uncertainty the only reason for holding money instead of goods will be to facilitate transactions, and the marginal rate of substitution of money for goods will represent the ratio between the net advantage of holding money for this purpose and of holding the goods concerned. If the marginal rate of substitution of goods for money is greater than their current price, holders of money will buy goods and the general price-level will rise, and vice versa.

There is no reason why such price changes should not be of a perfectly general character. There is nothing in the disequilibrium between goods and money which would alter the marginal rates of substitution between one good and another or between goods and debts. We should, therefore, expect that a change in the quantity of money would lead to equal proportionate changes in prices distributed over all goods—we are back in a quantity-theory world.

We now suppose that, for the first time, uncertainty about the future of the marginal rate of time preference makes its ugly appearance. We then have the emergence of a precautionary and a speculative motive for holding money; the marginal rate of substitution of money for goods will rise, and we may expect prices in general to fall (assuming, of course, that the quantity of money remains fixed). But this time the change in prices will not be spread uniformly over all goods. For a future change in the marginal rate of time preference would affect most the value of very durable capital and consumption goods and of very long-term debts. These will thus become more risky, and their price may be expected to fall relatively to less durable goods and shorter-term debts. This does not conflict with the fact that long-term rates may be lower than short-term ones if the latter are generally believed to be exceptionally high. This means, of course, that the prospective benefits from durable or long-dated assets are being discounted at a higher rate than previously. Our market rate of interest (or rather the complex of market rates on various types of asset) appears as the sum of our original marginal rate of time preference plus a risk premium.

Unfortunately it seems impossible to separate the two elements, but that is no reason to doubt their existence. The position is analogous to that of the price of industrial equities, which depends on an estimate of the future earnings of the company plus a rather different type of risk premium. No one can say the exact amount of this risk premium, but no one doubts that it exists.

The significance of interest in conditions of market equilibrium is simply that it is the rate of discount of future benefits which is necessary in order to induce the public to hold just that amount of durable goods of debts which happens to be in existence.

B. Having thus analysed the necessary conditions for equilibrium at a particular moment, we now turn our attention to the condition of stationary equilibrium over a period of time in which the rate of production of all goods remains constant with zero net investment. The conditions required for such equilibrium can be stated quite simply. First, it is necessary that the marginal rates of substitution between all possible combinations of goods shall be equal to the ratio of their costs of production. If, for any pair of goods A and B the marginal rate of substitution of A for B was greater than the ratio of their production costs, more of A and less of B would be produced and the marginal rate of substitution of A for B would be brought down until equilibrium was attained and vice versa.

Secondly, it is necessary that the rate of discount implicit in these marginal rates of substitution, so far as they affect durable assets, shall be equal to the market rate of interest; otherwise net investment or disinvestment, net saving, dis-saving, will take place. Thus in conditions of static equilibrium the rate of interest is that rate of discount of future benefits which is necessary in order to induce the public neither to add to nor to reduce its stock of durable goods and debts.

C. The possibility of net saving or dis-saving brings us to the consideration of a dynamic economy in which saving or dis-saving is going on, and brings us back to Keynes's fundamental criticism of the 'classical' theory of interest. For changes in the rate of saving or dis-saving imply changes in the level of income, and changes in the level of income will affect the supply of savings. It was for precisely this reason that Keynes maintained that the classical theory of interest was indeterminate. We have, therefore, to consider the circumstances in which changes of income can arise, and consider their significance when we find them.

Here, again, we must use a simplified model, but we need not go into the extreme simplicity of barter. Changes in income, according to the Keynesian analysis, arise from divergences between what savers plan to save out of a given income and what investors are prepared to invest. They can thus be traced ultimately to the fact that saving and investment result from separate decisions undertaken by separate individuals and from quite distinct motives. We shall, therefore, illustrate this in an extreme and highly unrealistic form by supposing a society divided into entrepreneurs who do all the investment and employees who do all the saving, both, of course, being also consumers. We can summarize the relations of each group as follows:

Entrepreneurs buy services from employees and sell consumer goods and debts to them. They also deal among themselves in capital goods and consumer goods. *Employees* receive wages from entrepreneurs and buy goods and debts from them, but do not deal among themselves.

Suppose that this society has a money economy, that there is no uncertainty about the future of the rate of time preference, so that the market rate of interest and the marginal rate of time preference will coincide, and, for the moment, that it is on static equilibrium with zero net investment—hence it follows that the prospective yield of capital assets must be equal to the marginal rate of time preference. Let the system have a net income of 1,000 money units a year, of which 800 go to employees and 200 to entrepreneurs. The money flow of such a system will be:

Entrepreneurs to employees (for services)	800
Employees to entrepreneurs (for consumer goods)	800
Internal transactions of entrepreneurs	200

and the total demand for consumer goods would be 1,000 and for capital goods zero, all quite consistent with equilibrium.

Now suppose that the marginal rate of time preference were to fall, and that everybody was convinced that the new rate had come to stay. There would thus be no uncertainty, and no incentive to hold money except for the transactions motive *however low the marginal rate of time preference might drop.*

The immediate effect of our reduction in the marginal rate of time preference would be to divert demand from consumer goods to debts. The price of debts would rise and entrepreneurs would be tempted to borrow more. Let us suppose that in the new situation 100 money units a year were saved. The money flow of the system would then be:

Entrepreneurs to employees:	
For services	800
Employees to entrepreneurs:	
For goods	700
For debts	100
Internal transactions of entrepreneurs:	
For consumer goods	200
For capital goods	100

Total demand would be 900 for consumer goods and 100 for capital goods, still maintaining our income of 1,000 units a year.

In the worst possible circumstances, if entrepreneurs could find no use for additional capital even on the more favourable terms now offered, the supply of debts would remain the same and their price would rise to the point at which employees' demand was choked off and they continued to spend as much as before on goods. Since no one has any reason for trying to increase his stock of money, any income which is diverted from consumer goods must swell the demand for capital goods and we can have no deficiency in total demand. Again, for the moment, we are comfortably tucked away in a Ricardian world.

Conversely, so long as entrepreneurs can only buy capital goods by selling debts to savers, they can only invest what the savers are prepared to save, and the system can no more suffer from inflation than it can from unemployment.

We now again admit uncertainty, with the result that the marginal rate of substitution of money for debts rises, the price of debts falls and their yield rises, and employees set aside part of their savings to increase their stock of money. The money flow now looks very different. Suppose that out of 100 units saved 50 go to increase employees' money stocks. Then we have:

Entrepreneurs to employees: For services	800
Employees to entrepreneurs: For goods	700
For debts	50
Internal transactions of entrepreneurs: For consumer goods	200
For capital goods	50

Total demand will now be only 950 units, and entrepreneurs will have run down cash balances and accumulated unwanted stocks of 50 units. They will try to recover their cash balances, by curtailing production and selling off stocks, and so we are in for a fall in income. But the increased propensity to save led to a fall in income only when it was associated with an increased propensity to hoard, and the propensity to hoard came in under the coat-tails of uncertainty as to the future rate of time preference. Conversely, an inflationary situation can develop when an increase in the quantity of money drives the marginal rate of substitution of money for debts below the marginal rate of time preference, or when the marginal rate of substitution of money for debts fails to rise in response to a rise in the profitability of investment.

It thus appears that the rate of interest is firmly based upon the marginal rate of time preference, but that uncertainty as to the future of the marginal rate of time preference usually keeps the market rate of interest above that rate.

It is thus uncertainty as to the future of the marginal rate of time preference that complicates the straightforward relationship between money and prices supposed in the quantity theory. It is this uncertainty that gives money its peculiar significance in economic dynamics, since money gives both the opportunity of escaping this uncertainty and the hope of turning it to one's own advantage. And it is this uncertainty which, through the propensity to hoard, is responsible for changes in the total level of effective demand and, therefore, of income.

UNIVERSITY COLLEGE
OF SWANSEA.

STATIC MODELS AND CURRENT PROBLEMS IN INTERNATIONAL ECONOMICS

By T. BALOGH

SIMPLIFYING assumptions, and the construction on their basis of schematic models of the system to be investigated, as an aid to analysis, is a necessary procedure in all sciences. Economics, beset with all the complications in the nature of social studies, is certainly no exception to this rule. But authors availing themselves of such tools must remain conscious of exactly what they are doing, and must, in particular, be bound by the limitations of their models. To take an example from the natural sciences, it would be quite proper to construct a 'model' by assuming that light acts as if it consisted of particles of colloid crystals. What might perhaps be considered as somewhat unscientific is to refuse to investigate closely whether this hypothesis can be verified, or claim (supposing that it were mathematically a 'simple' case) that in order to save the determinateness of the solution, or the mathematical beauty of the structure elaborated, we should continue to work and prescribe further experiments on the basis of the colloid model, the facts of life notwithstanding.

These observations may seem trite. Yet how much of economic doctrine violates these simple canons! How much of the current disputes originates in the fact that protagonists conduct their case on different levels of abstraction while trying to advise on the same practical matter!

Of late the 'new look' in economic models, attempts to construct economic models on the basis of statistical correlations and to predict future trends as a guide to policy, has come in for severe censure. The fact that interest in such explorations awakened mainly during the war, when the material, if available at all, was often unsuitable for the elaborate use made of it, has undoubtedly justified much of the criticism. Few objections were raised, however, to other, less explicit, less easily disproved, literary 'models' which are utilized by 'pure' theory. Yet it might seem that, however simple and wrong some of the econometric models have proved, they have infinite advantages over the beautiful gossamer structures of pure theory. Their assumptions are disclosed and the manipulations are patent. Moreover, they frankly try to build on the basis of scientific principles and are put forward for the purpose of testing. Even if the first few hundred attempts fail, they will have performed the useful function of elucidating the complicated interrelations in social development.

In contrast there is little attempt at even a perfunctory verification of the assumptions of the 'pure' theoretical models. Indeed, a certain pride is shown in their uselessness if they are subject to attack; though this pride is usually not manifested when, in the absence of attacks, they are used as a basis of policy recommendations.

The present paper will discuss a model of this kind which is used by Professor Samuelson in his brilliant article on 'International Trade and Equalization of Factor Prices' in the *Economic Journal* (June 1948). Professor Samuelson wishes to examine closely the theorem that commodity trade alone can only achieve a partial equalization of factor-rewards. He develops his argument on the basis of 'the simplifying assumptions most suited to the Ohlin analysis'.[1] Two perfectly immobile factors of production are assumed to be distributed in unequal proportions between two regions. It is further assumed that

(*a*) the same production functions exist in both countries, i.e. that the productive factors are identical and work in identical surroundings;
(*b*) the productivity of the region is not affected by the scale of operations, i.e. there are constant returns;
(*c*) the productivity of factors is only affected by changing their proportions, i.e. they are completely homogeneous.

Professor Samuelson contends that if these assumptions are granted, Professor Ohlin has not gone nearly far enough in claiming that 'free mobility of commodities in international trade can serve as a partial substitute for factor mobility, and will lead to a partial equalization of relative (and absolute) factor prices'. He demonstrates that the equalization would be complete.

The proof of the theorem consists in comparing the two production-possibility curves of the two regions. These are assumed to be continuous. Equilibrium will be reached when the point of common slope on the two curves is reached. Then ratio curves are constructed for the two industries, one axis being the ratio of the quantities of the two factors used, the other of their real incomes or physical marginal productivities. These curves are again continuous and *identical* for *both* countries, as the technical production functions are identical. They are contrasted with each other, and with a curve constructed showing the relation of the ratio of factor rewards to the ratio of the prices of the two products. At the point of equilibrium in both countries identical ratios are achieved: the proof is unanswerable.

---

[1] Professor Samuelson might be thought slightly less than fair to Professor Ohlin in so far as the latter recognizes the importance of increasing returns (chap. iii, pp. 52–8). Unfortunately there is at no point of Professor Ohlin's statement a clear summary of his assumptions and a large part of the discussion is conducted on assumptions which exclude increasing returns as 'not compatible with long-run equilibrium'. In some way the limitations of Ohlin's method are shown by his belief that foreign trade will 'reduce the disadvantages' caused by indivisibility. He completely fails to see that his treatment does not differentiate between labour and other factors of production from a welfare point of view, and disregards the grave danger of 'free trade' tending to perpetuate the disadvantageous original distribution of factor endowments in cases of increasing returns. The supply of child labour, for example, cannot be taken as a datum when judging economic policy. Moreover, as we shall demonstrate, endowment with 'capital' is, at least partly, a function of the policies pursued.

Professor Samuelson sums up his conclusions as follows:

> '(1) So long as there is partial specialisation, with each country producing something of both goods, factor prices will be equalised, absolutely and relatively, by free international trade.
>
> '(2) Unless initial factor endowments are too unequal, commodity mobility will always be a perfect substitute for factor mobility.
>
> '(3) Regardless of initial factor endowment, even if factors were mobile they would, at worst, have to migrate only up to a certain degree, after which commodity mobility would be sufficient for full price equalisation.
>
> '(4) To the extent that commodity movements are effective substitutes for factor movements, world productivity is, in a certain sense, optimal; but at the same time, the imputed real returns of labor in one country and of land in the other will necessarily be lower, not only relatively but also absolutely, than under autarky.'

He adds specifically:

> 'Factor-price equalisation is not only possible and probable, but in a wide variety of circumstances it is inevitable.'

These are indeed surprising theorems.

Professor Samuelson concedes that he has proved too much and turns back accordingly to the assumptions on which his analysis has been conducted, in order to explain by their imperfections the discrepancy between his logical conclusions and the very different reality. He finds the most serious imperfections in the assumptions which underlie the limitations of 'factor-proportions analysis'. It is the great merit of his article that it has brought these to light by formulating them systematically and clearly.

Yet perhaps Professor Samuelson is open to the charge which he brings against Professor Ohlin of not pushing his reasoning far enough. His article is tantamount to a damaging *reductio ad absurdum* of the assumptions which underlie modern developments of international trade theory: and remembering some of his earlier exploits, I thought at first that this might have been his object in writing it. But this interpretation will not survive a careful scrutiny. Professor Samuelson, it is manifest, retains considerable tenderness for the assumptions which lead to such strange conclusions; nor do the conclusions seem so strange to him as they do to others. Indeed he treats them as sufficiently reasonable to provide the basis for a practical moral as to the dubious wisdom of large-scale emigration from Great Britain. This moral is sensible enough in itself; but it can derive no real support from an analysis based on assumptions as artificial as those of this model.

The defects of these assumptions are much greater and more fundamental than Professor Samuelson brings out.

1. The problem of defining identical productive factors in the two regions involves even greater difficulties than Professor Samuelson suggests:

homogeneity in this sense would rule out the most important observable cause of increasing cost: the less than perfect substitutability for one another of different doses of the same 'factor'. If this is admitted the assumption that identical production functions will exist in the two regions must be dropped. It is so highly improbable that precisely the same pattern of imperfections in substitutability should be encountered in both regions that this contingency can be dismissed. With it goes the proof and the theorem.[1]

2. The assumption of constant returns to scale rules out most foreign trade problems historically encountered. An attempt to 'assimilate' indivisibilities by regarding each of the indivisible units of a factor as a separate factor would—like the assumption of homogeneity of factors and its consequences—reduce the 'explanation' to a farce. For once we admit that indivisibilities within broad factor categories exist, then the historical size of the market becomes of paramount importance for the *further* development of the economy. Once an advantage of this kind is captured by a *country* it will have an automatic tendency to grow cumulatively. Even those checks which are held to prevent a too easy and complete victory of a *firm* which has snatched a superior position from its rivals—the growing difficulty of effective organization and supervision beyond a certain point, and the increasing rigidity in entrepreneurial decisions—will not operate in the comparative position between countries even to the limited extent they are observable in industry.

Such indivisibilities might be internal to firms in the regions. The effective limit of these will depend on the size of markets. The uniformity of tastes and customs which characterize a community will have a decisive influence in determining the possibility of utilizing them. Nor will there be in *international* relations the same mitigating elements of cost, i.e. sales cost, which are encountered by *firms* when they try to establish a special market for their own products. In international relations these 'sales' costs are mostly borne by the community: the common schooling and tradition creates the cohesion of the market. These costs are not lower for smaller and poorer than for richer communities. Moreover, the 'prestige' value of 'tastes' and habits in the leading community will represent a protection in the home and a preference in foreign markets. The imperfections of the international market are therefore much stronger and wider in scope than those of single industries usually treated. It is a pity that Professor

---

[1] Moreover, it is illicit to assume that the opening of commodity trade will not affect the categories of productive factors, stimulating or obstructing their development or change. If, for example, in an underdeveloped area no possibility exists to acquire industrial skill and thus total specialization seems indicated by the respective 'curves', this condition will be perpetuated by trade with a more developed country, however advantageous it would be for that area to acquire skill.

Allyn Young's suggestive article[1] had so small an effect in its application to the theory of international trade. The problem of international trade and its effects on the evolution of individual member countries of a world economy—much more even than the problem of economic evolution in a given community—involves the interaction of societies on widely *differing levels of development*. This is a totally different situation from the mutually beneficial exchange of commodities between traders of equal standing, subject to similar dynamic processes in their potential development, to which the considerations applicable to simple barter could be extended. Open relations viewed from this angle might be held to jeopardize the eventual achievement of equal opportunities by the more primitive or smaller community for the immediate gain from trade, disregarding the possibility of achieving a similar position later.[2]

The story only begins here. The size and wealth of the community itself create increasing returns over a wide range of activity. The rise of specialized firms, the extension and cheapening of services might be mentioned.[3] The growth of income thus engendered results in a more than proportional increase in capital, especially of the determinant 'own' capital of firms, as investment habits become adapted and shares can be placed. Far more important, however, is the cumulative increase of opportunity and also of the facilities to train or create special, indivisible 'factors of production'. The growing pull of the dominant economy for the best productive factors, especially entrepreneurial and technical ability of *other* regions, is hardly less important. Thus the discrepancy in indivisible factor endowment is further intensified. To neglect these essential features of economic growth is to neglect the most significant factors in the development of first England, then Germany, and, in our days, of the United States, and thus the real determinants of the pattern of international trade.

It has been objected that the size and wealth at any rate of the greater western countries is large enough to permit the organization of optimal units in almost any branch of industry. Moreover, by engaging in international trade it is said that the limits of *national* markets can be overcome.[4]

---

[1] 'Increasing Returns and Economic Progress', *Economic Journal*, Dec. 1928.

[2] Graphically this would correspond to cost curves with several inflexion points. The 'backward' community cannot overcome the barrier before the second, &c., range showing increasing returns is reached and has to trade on the basis of the least favourable relationship, and so on. Education takes time. Social services are equally important.

[3] Including the supply of capital. Risk is reduced by low gearing based on a well-developed capital market. The financial strength of the government plays an equally important role. See also additional note at end.

[4] In this connexion it is sometimes pointed out that Britain is the third or fourth largest manufacturing nation, with population eighth in size in the world. But the problem is not whether there are smaller countries but that the two wars enormously enhanced the power of two *supra-national*, continental entities which differ in basic economic character from the rest of *national* economic organizations.

These objections also fall into an illicit simplification of the conditions under which international trading takes place, and particularly the element of differential risk attached to foreign trade. Secondly, they disregard the importance of historical development in the evolution of patterns of industrial organization which cannot be explained by mechanical reference to maximization of profit.

In modern conditions the existence of a large market possessing certain definite tastes and habits confers a substantial advantage to producers in that area (especially if it is protected). There is first the possibility of unfavourable regulations of sovereign governments over which producers outside the area have no direct influence, which renders foreign trade more hazardous than domestic trade. Secondly, adaptation to varying habits will necessarily lose the exporter some advantages enjoyed by the producers working mainly for a standard market. This is the sound core of the argument against the proposals to introduce mass-production methods into British export industries, though it is usually overdone, in so far as it neglects the influence lower prices might have in overcoming 'tastes'.

It would be possible by imposing standardization and concentration to overcome some of the disabilities attached to smaller markets. But it would be rash to expect that entrepreneurs acting on their own initiative would be inclined to such reorganization. First and foremost common prudence counsels that they should distribute and limit their risk among a variety of products. Secondly, they are committed to a certain pattern and will be unwilling to start a mortal combat to eliminate their competitors. The so-called kinked demand curve is merely a picture of a frame of mind, yet it is a reality with which we must reckon. Thirdly, they will be afraid of the political repercussions of an attempt to reduce to a very few the number of firms in existence. These 'imperfections' are so important that no realistic discussion of foreign trade problems should disregard them. Their existence certainly rules out in practice such reorganization of industry as would permit a smaller country to equal the productivity of larger markets.[1] If State ownership of the industries is to be avoided, and it must be remembered that the fate of most heavy and mechanical industries is at stake, the creation of a large market is the only way in which differences of productivity can be overcome. But even publicly owned export enterprise (i.e. one not subject to bankruptcy and annihilation) might be difficult to start in the absence of some certainty of success after a transition period which will in any case be difficult.

These considerations would seem to indicate that models which *a limine* exclude the possibility of increasing returns must be ruled out for the illumination of problems of international trade.

---

[1] 'Underdeveloped' countries are at some advantage in escaping this cul-de-sac; for a new industry would be either organized on the most modern lines or not at all (e.g. the industrialization of Canada). But older industrial *exporters* are not in this position.

3. A further and connected problem arises in respect to the treatment of capital as a productive factor. In a static model it is quite justifiable to assume that quantity of capital is 'given'. The production pattern of the economy is then also 'given'. If then commodity trade takes place between these two static communities—and if we disregard all that has been discussed in the preceding paragraphs—some sense might be attached to problems of 'optimum' distribution of factors.

Unfortunately for this view the supply of capital is to a large extent a function of investment possibilities. This includes the degree of uncertainty which itself is not unrelated to the scope of foreign trade. In a less-developed country faced with potential competition of a more developed it will be influenced by that threat.[1] It also includes the cumulative processes enforced in the poorer country by the readjustment, which it will be less capable of sustaining unscathed. The eventual position of equilibrium will then be shifted to its disadvantage. The influence of the other factors making for increasing returns is thus further strengthened. The greater the accumulation of capital the easier further accumulation, the more smooth the equation of supply and demand through appropriate institutions, the existence of which depends on a certain availability of capital.

Indeed, if we take all these highly potent forces into account, one must wonder why trees do not grow into the sky, killing by their shadow all other plants which were left behind. The tendency to sharper cyclical reactions, the financial defaults on foreign debt which effect a potent redistribution of income, the pressure of protection, are perhaps the main headings under which an explanation has to be sought.

4. This leads us to the last point.

The concept of 'equilibrium' is highly unrealistic when employed to judge economic development in any one country, i.e. in comparing positions *through time.* And it loses all pretence at realism when it comes to an analysis of the *interaction of societies* which differ in economic development and social habits. Foreign trade then represents the sudden bringing together and continued coexistence of economies at different points of development through time. New wants and new products disrupt the structure of indifference curves. Complementary sets of wants, determined by social background, the transition between which is non-continuous, render impossible analysis by simple yet highly abstract hypotheses about the general behaviour of consumers. Nor is there any determinate *a priori* answer to be found about the likely shape of the path towards equilibrium.

---

[1] More positively these considerations necessitate a complete revision of the concept of 'infant' industry and its extension to whole countries and regions. Cf. my paper on 'Some Theoretical Problems of Post-war Foreign Investment Policy', *Oxford Economic Papers*, No. 7, Mar. 1945.

If the optimum of general theory is scarcely applicable to the analysis of consumers' wants in a developing economy, and its relations to other systems, the analysis of production-possibilities is even less legitimate, even if we did not encounter those difficulties of a conceptional character which have been discussed above. They all imply an infinite number of production combinations, the transition from one to another of which is smooth and costless. Nothing is farther from the truth. The linkage in production-lines is intimate, and in the overwhelming majority of instances the private cost of starting any new line of development is dependent on the parallel increase in other lines (if for no other reason because of the uncertainty of markets).[1] Social and private costs differ widely.

Nor is this all. The combination of productive factors, especially of capital, is not possible in infinite variety. Only a limited number of definite patterns are technically feasible. The transformation of a combination into another takes time and gives rise to formidable social costs which cannot be imputed through the market mechanism. If, however, the indivisibilities are large relative to marginal changes, the whole approach which tries to concentrate on the analysis of the latter is vitiated. We do not know the possible points in the production-possibility map which can exist, and we cannot say anything about the slopes, i.e. the ratios, at these unknown points. The stability of these relations and the path of readjustment are equally in doubt. The analysis yields no determinate answer.

Professor Samuelson's brilliant mathematical feat should earn him the deep gratitude of economists. By clearly setting out the basis of the most modern variant of the 'pure' theory of international trade he has at last cleared the way for a new dynamic approach to this essentially dynamic problem.[2]

---

[1] Cf. my paper on 'Foreign Investment', *Oxford Economic Papers*, No. 7, Mar. 1945, p. 106, on the differences between current private and future social costs.

[2] After the completion of this paper Professor Samuelson published a further article on his theorem (*Economic Journal*, 1949). It further strengthens the mathematical proof of the 'theorem' discussed in this paper without investigating the implications of its assumptions. Thus there is no cause for modifying or amplifying the present analysis: Professor Samuelson seems to have been even more strongly captivated by mathematical beauty than is assumed in this paper.

Additional note to p. 195 *n.* 3. All this culminates in the power to grant foreign loans, which (even if they are not 'tied') represent the most important influence of all in favour of the dominant economy. The 'pure' theory of international trade altogether excludes this problem by its stringent assumptions—on which, however, its more important normative conclusions depend.

BALLIOL COLLEGE,
OXFORD.

THE THEORY OF COMPARATIVE COSTS RECONSIDERED

By G. M. MEIER

REAWAKENED interest in the pure theory of international trade[1] and the development of the 'new' welfare economics make it desirable to examine once again the doctrine of comparative costs. Consequently the purpose of this paper is to synthesize various interpretations of the principle of comparative costs, to emphasize certain elements which are frequently neglected, and to suggest a reformulation of the doctrine's welfare conclusions.

Section 1 examines the relevancy of comparative costs to the general theory of value. It will be contended that the doctrine is as applicable to a closed economy as it is to an open economy. Though the doctrine may be fundamentally relevant to the general theory of value, nevertheless its traditional form is not realistic. Section 2 considers this lack of realism by summarizing the Ricardian analysis of comparative costs in its classical and refined neo-classical forms and by considering the merits of various criticisms levied against it. It will be maintained that advances in the general theory of value have left a gap between the traditional comparative cost analysis and modern value theory. A restatement is therefore necessary. Section 3 attempts such a restatement, using indifference-curve techniques and general equilibrium theory. Grounds will be submitted for considering this presentation to be superior in some respects to the traditional view. However, as regards the gains from trade, it will be maintained that this restatement is also limited. Consequently, the final section suggests that a partial welfare analysis based upon the surplus concept may be more appropriate for analysing the relation between international trade and economic welfare.

1

At the outset we may establish the relevancy of the comparative cost doctrine to the general theory of value. Both internal and international trade theories must explain the value relation among commodities which satisfy different sorts of demands and which are made under different conditions of production. The theory of market exchange provides the proximate answer to this problem. But just as there is some type of subjective valuation behind this theory on the demand side, so is there an ultimate choice of specialization on the supply side. Since comparative costs explain the basis of specialization, the doctrine is merely an elementary premiss of the theory of exchange—whether the exchange be among individuals, regions, or countries. In fact, to show the logic of the doctrine, one may substitute the word 'individual' for 'country'. Why do individuals trade with each other instead of each one producing for his own requirements ? Why does division of labour increase the total efficiency of production ? The answers to these questions are usually presented in the first pages of an economics text-book. Suffice it to recall that the explanation involves the facts of 'varying abilities' and the 'advantages of specialization'.

---

[1] See Joan Robinson, 'The Pure Theory of International Trade', *Review of Economic Studies*, xiv (2), pp. 98–112; P. A. Samuelson, 'International Trade and the Equalisation of Factor Prices', *Economic Journal*, lviii (June 1948), pp. 163–85.

The simplicity of these questions, however, should not obscure their fundamental relation to the theory of exchange and, in turn, to the theory of value. Not only does the theory of comparative costs establish the basis for specialization, but it also shows under what conditions trade will be profitable to all concerned. It thus has two roots—one in value theory and the other in welfare theory.

This is evident when it is realized how closely related are comparative costs, the Robinsonian theory of rent, and the Marshallian surplus concept.[1] Given a country with a certain block of resources, how are these resources to be allocated among different commodities ? This is the problem with which the theory of comparative costs is concerned. When this problem is viewed as one involving alternative costs, the comparative cost doctrine becomes related to the Robinsonian theory of rent.

Furthermore, the relevancy of comparative costs to rent theory and to general value theory may also be noted in the familiar 'cost ladder' of an industry's supply curve.[2] In such a curve the intra-marginal firms have a comparative advantage in that particular industry; the firm with the lowest minimum average cost has the greatest comparative advantage and earns the largest rent in the Robinsonian sense.

It is also significant that in early private papers and in the first edition of his *Principles* Marshall referred to consumer's surplus as consumer's rent.[3] The connexion between comparative costs, rent theory, and the Marshallian surplus concept is thus a direct one.

However, the place of comparative costs in the general theory of value is most readily seen when the Ricardian analysis of comparative costs is compared with the Walrasian theory of general equilibrium. For it then becomes apparent that the Ricardian theory and the Walrasian system are but limiting cases in the general theory of value. The Ricardian analysis recognizes only one factor in a country (Portuguese labour or English labour) and only two products. As a special case of the more general theory, the Ricardian analysis emphasizes complete specialization and determines whether it is cheaper for one country or the other to produce a given commodity once the terms of trade (the Walrasian factor–price ratio) are known.

---

[1] I am indebted to some lecture comments by Mr. J. R. Hicks for the line of reasoning in the next four paragraphs.

[2] Cf. K. Boulding, *Economic Analysis* (1941), p. 438.

[3] Alfred Marshall, *Principles of Economics* (1890), p. 175.

In contrast, the Walrasian system recognizes several factors of production, variable supply of factors, and introduces demand to solve the problem of multiple exchange with more than one factor. The Ricardian analysis, furthermore, cannot be translated into the form of equations but only inequalities.

But if the comparative cost doctrine is so basic to the theory of exchange and has so many analogues in general value theory, why, then, has it not been used more in general value theory?

Of course, if one maintains that the doctrine must be presented along the classical lines of a labour theory of value, then aversion to this methodology may alone be sufficient to prevent its use. But the reason is probably more substantial. It is to be found in the different approaches which neo-classical literature has used when considering a closed and an open economy. To the neo-classicists the theory of value in a closed economy has been basically of a partial equilibrium character, whereas their theory of international trade resembles a general equilibrium theory. In the latter approach recognition of comparative costs is necessary for an analysis of the system of prices for the economy as a whole. Moreover, the theory of domestic trade is usually stated in terms of comparative prices, so that little is directly said about comparative costs. Yet the relation between prices and opportunity costs is an intimate one, and the result will usually be the same regardless of which method is used. It is hoped that this essential relationship will be evident after the comparative cost doctrine is restated in the modern form which is more harmonious with general value theory.

However, before offering this restatement let us first consider the realism of the doctrine in its traditional form.

2

The well-known Ricardian tables of the Portugal–England type answer two questions: (1) What commodities will a country export? (2) What are the limits within which exports will exchange for imports?

The Ricardian answer to the former is that a country will export those commodities in which its labour costs of production are relatively lowest, ignoring transportation costs. In modern terminology the Ricardian model is concerned with total conditions as illustrated by its emphasis on complete specialization. This emphasis is more valid than is commonly supposed, as will be seen in a subsequent section.

The Ricardian answer to the second question is that relative costs govern the rates of exchange between commodities and establish as limits in foreign trade the terms of exchange of the commodities produced in each country. Moreover, it is maintained that if trade is free each country will specialize according to its comparative advantage, and a mutual benefit will result for all participating countries.

The assumptions underlying the Ricardian analysis are numerous.[1] Restrictive as these assumptions may appear, nevertheless the successors to Ricardo have done remarkably well in relaxing many of them and in generalizing the analysis. To realize this we need only recall the contributions of J. S. Mill, Cairnes, Edgeworth, Marshall, and Taussig.[2] For our immediate purposes it will be sufficient merely to observe the forms in which the doctrine was left by Taussig and Marshall.

Taussig's general exposition is certainly within the classical tradition but is more general in its application than is the Ricardian formulation. For Taussig introduces differences in wage rates, money prices, capital charges, variable costs, numerous commodities, numerous countries, and transportation costs. Nevertheless, the end-result of his analysis is quite similar to the pure Ricardian presentation. For in essence Taussig maintains that differences between comparative money cost and comparative labour-quantity cost are so slight that comparative costs in terms of labour-quantity costs may still be considered valid. This reasoning is based on the 'hierarchy principle' by which Taussig contends that the hierarchy of labour in various countries is likely to be the same in both rank and percentage differences. The same principle is also applied to the ratio of capital costs to total costs in different industries in various countries.[3]

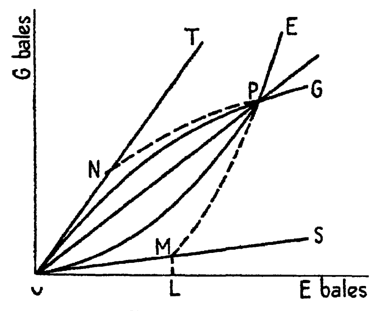

DIAGRAM I

This conclusion is the relevant one for our present purposes. We should also recall, however, Marshall's geometrical version of comparative costs, for it is a logical antecedent of the modern restatement to be discussed shortly.

The Marshallian reciprocal demand analysis is a general equilibrium type of analysis, stemming from J. S. Mill's principle of 'the equation of reciprocal demand'. Marshall measures all the export-goods of a country in terms of a common unit—the 'representative bale'. Diagram I illustrates the Marshallian analysis for 'normal' conditions of international trade between two countries (*E* and *G*), the actual equilibrium position of trade being at *P*, with the slope of *OP* representing the terms of trade.

---

[1] Cf. Bertil Ohlin, *Interregional and International Trade* (1935), Appendix III.
[2] See Jacob Viner, *Studies in the Theory of International Trade* (1937), pp. 493–516.
[3] F. W. Taussig, *International Trade* (1927), pp. 51–2, 67–8.

Following Edgeworth, but including Viner's corrections of the analysis,[1] we may show the relationship between comparative costs and the Marshallian reciprocal demand curves. If in Diagram I we now measure German linen on the *Y*-axis and English cloth on the *X*-axis, the lines *OS* and *OT* represent, on the assumption of constant costs, the ratio of the cost of production of a unit of linen to that of a unit of cloth for England and Germany, respectively. *OS* thus represents the terms on which England could obtain linen, and *OT* the terms on which Germany could obtain cloth in the absence of trade. However, if in the absence of trade the amount of linen which would be produced and consumed in England is equal to *LM*, the English export supply curve of cloth in terms of linen will be *OME* rather than merely *OE*. The equilibrium terms of trade must fall between these two lines in order to allow England to obtain linen and Germany to obtain cloth more cheaply than they could without trade.

It is also essential to recognize that the Marshallian analysis introduces real costs in an explicit form. As long as we are willing to retain the labour theory of value a 'representative bale' has some economic meaning, and we may follow Marshall's analysis. But once we assume that there are many different qualities of labour, that other factors of production besides labour are available, and that many factors are specific, then we must discard the labour theory of value. Once we recognize these other assumptions, then a study of modern value theory will show that the labour theory is valid only as a special case of the general theory when the only factor is homogeneous labour. Indeed, if we limit the doctrine of comparative costs to its Ricardian and neo-classical forms a wide gap appears between the theory of international value and the theory of value in a closed economy. For in using a labour theory of value to explain comparative costs one has to assume among other things that the basis for calculations is the cost at the production margin, thus eliminating rent; that various categories of labour can be reduced to a common denominator; and that capital and labour are used in the same proportions in the production of all commodities. Surely these assumptions are neither realistic nor limited to the minimum number necessary for analysis.

General value theory has recognized this and has removed these assumptions by emphasizing valuation as a problem of choice and by using the concept of displaced alternatives. Moreover, the modern theory of value is readily translatable into a general equilibrium system. This widens the gap even more. For while the Marshallian analysis is a type of general equilibrium analysis, nevertheless it cannot be considered as refined as present general equilibrium theories. On the contrary, it is far less general and much more simplified, being inadequate for a solution of the perfectly general case in which there are any number of individuals, countries, firms, products, and factors.

---

[1] Viner, op. cit., pp. 547–8.

If general value theory has departed so sharply from a labour theory of value, why then should the theory of comparative costs be stated continuously in terms of that abandoned theory?

Viner maintains that the labour theory is essential to demonstrate the very *raison d'être* of the comparative cost doctrine. He insists that the doctrine was developed primarily to show the gains from trade. To show the gains from trade it is necessary to include real costs. To include real costs it is convenient to use a labour theory of value. Therefore the doctrine is justly stated in terms of that theory, even though it may have been abandoned in general value theory.

Significant as this argument is, nevertheless I believe its validity may be questioned. Firstly, it is by no means clear that the original Ricardian analysis actually was concerned with real costs. Time and again in his *Studies* Viner states that in the Ricardian doctrine cost 'meant real cost'.[1] To Viner it is definite that the classical writers were concerned with the subjective costs or disutilities, and that by real costs they meant all subjective costs directly associated with production. The implication clearly is that writers who consider 'efforts and sacrifices' are in the Ricardian tradition, whereas 'modern' theories which do not consider these 'efforts and sacrifices' are deviating sharply.

I think this is a view which is commonly held. However, Shove and Myint have recently emphasized that real costs do not appear in the Ricardian theory. An especially relevant paragraph in Shove's article deserves quotation here:

> '[In Marshall's *Principles*] the "supply price" of a commodity now becomes the sum of the prices which have to be paid to "call forth" the "efforts and sacrifices" which are required for making it and constitute its real costs of production—a quite un-Ricardian notion. For Ricardo, labour is not a "disutility", but the productive force available to the community, the stuff, so to speak, by means of which commodities are made, and the cost of a thing is the quantity of this force or stuff, together with the quantity of capital, absorbed in its production, not the effort and sacrifice entailed in providing it. . . . With the emergence of the psychological conception of "real cost" we enter not merely a different world from Ricardo's, but a different universe.'[2]

Thus, if this interpretation be true, real costs appear not in Ricardo but in the neo-classical presentation which Marshall epitomizes.

---

[1] Viner, op. cit., pp. 483, 488, 490–1.

[2] G. F. Shove, 'The Place of Marshall's *Principles* in the Development of Economic Theory', *Economic Journal*, lii (1942), pp. 306–7. Also see Hla Myint, *Theories of Welfare Economics* (1948), p. 26.

Secondly, many neo-classical theorists assumed in their analyses of comparative costs that money costs are proportional to real costs. From Senior and Cairnes on, writers frequently recognized besides the 'irksomeness' of labour other costs in the form of time-preference and risk. Consequently the unit of real cost cannot be in terms of hours of labour, and resort must be made to the assumption that market prices may indicate all these real costs through a proportionality factor. This requires the further assumptions that different factors are used in the same proportions in different industries; that, as Marshall stated, money income remains constant in terms of effort; that money costs and prices of all commodities are proportional to their marginal disutilities of labour; and that at the margin there are equal earnings in alternative occupations. Moreover, land or natural resources cannot be included in any real cost sense. If one is willing to accept these restrictions, then the modern restatement may also consider real costs.

However, if we are too sensitive to these assumptions and omit any direct reference to real costs, we may still demonstrate the gains from trade with the aid of modern welfare economics, as the following sections show.

We are thus in a quite pleasant situation. We may include the Marshallian type of real costs or not—just as we wish. The final objective of demonstrating the gains from trade may be realized regardless of our choice. Moreover, if we do choose to ignore subjective real costs we shall actually be within the purest of the Ricardian tradition and should not be considered as iconoclasts by Professor Viner.

Let us now consider how we may restate the comparative cost doctrine without employing a labour theory of value and with more precision in respect to demand than simply Ricardo's vague statement that foreign trade 'will very powerfully contribute to increase the mass of commodities, and therefore the sum of enjoyments'.[1]

3

Our starting-point is a theory of value based upon opportunity costs. Opportunity costs do not explain why the different ratios of exchange should exist in the two countries concerned. But to understand this we need only refer to the Heckscher–Ohlin analysis, which establishes the reasons for comparative advantage by analysing the relative supplies of factors entering into the various production functions of the respective countries.

Following Samuelson,[2] let us take as given all the technological relations so that we may then have a production function in the implicit form: $\phi(x, y, a, b) = 0$, where x and y are the amounts of commodities produced per unit of time, and a and b are the amounts of productive services rendered per unit of time. Given the amounts of any three of the variables, the fourth is determined. If we now regard a and b as fixed in amount, the resulting relationship between x and y is the familiar production possibility curve MN as in Diagram II.

---

[1] David Ricardo, *Principles of Political Economy* (Gonner ed., 1891), p. 108.

[2] P. A. Samuelson, 'Welfare Economics and International Trade', *American Economic Review*, xxviii (June 1938), p. 262.

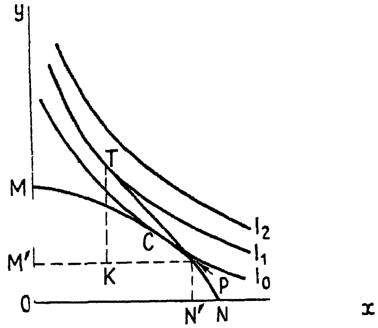

DIAGRAM II

It must be realized, however, that such a curve as MN shows the maximum amount of x and y that can be produced from a given quantity of the factors a and b. This follows from the fact that the curve has been derived from the Edgeworth–Bowley type of box diagram.[1] The locus of the points of tangency in such a diagram forms a 'generalized contract curve' which gives an infinity of optimal production positions. This is due to the condition that at each point of tangency the ratios of the marginal physical productivities of each factor are the same in each industry. The generalized contract curve is not the production possibility curve (henceforth referred to as p.p.c.), but by reading off the amounts of x and y along this curve and by translating to a plane such as in Diagram II we do get the p.p.c.

Certain properties of this curve should be observed. If the production function is homogeneous in the first degree the p.p.c. will be linear, corresponding to the Ricardian case of constant costs. If, however, production is under increasing costs, the curve will be concave downward; under decreasing costs, concave upward.

Moreover, the greater the degree of substitutability among the factors, the flatter will be the curve. In contrast, if most of the factors are specific to the production of either x or y, the curve will have a bulge at a certain point beyond which the specificity limits production. In the extreme case, if no factors could be transferred between x and y, the curve would be similar in shape to $M'PN'$ in Diagram II.

As to a consideration of time, just as the shape of the Marshallian supply curve varies according to the length of time which elapses, so too will the shape of the p.p.c. vary. The longer the run allowed for production to adapt itself, the flatter will be the curve.

Finally, it is essential to note that at a point on the p.p.c. showing the combination of x and y which is produced, the slope of the tangent to the curve at that point will be the price ratio. And this will be equal to the substitution ratio between x and y, which, in turn, will be equal to the ratio of the marginal money costs of x and y. Thus in Diagram II, at the point P, the slope of PT is the price ratio which equals the slope of the substitution curve which, in turn, equals the ratio of marginal costs. All these equalities follow from the conditions of pure competition.

---

[1] Cf. W. F. Stolper and P. A. Samuelson, 'Protection and Real Wages', *Review of Economic Studies*, ix (Nov. 1941), pp. 58–74.

We have thus already made considerable progress without any reference to a labour theory of value. We have considered different factors of production, have recognized the possibility of various kinds of production functions, and have shown that, given the demand, the relative prices of the commodities will be determined by their costs. Moreover, by introducing money costs directly into the analysis we have proceeded away from not only labour cost but also opportunity cost.

To complete this restatement of the comparative cost doctrine, we must now answer the same questions which the traditional presentation does: What commodities will a country export and import? And how can the gains from trade be demonstrated?

Consider the case of one country as represented in Diagram II. If the economy is a closed one, demand conditions may be such that production is at the point C. Now, however, external trade is introduced, and the terms of trade or external price ratio is given as the slope of PT. Production in the country will move to point P where the internal ratio of marginal costs becomes equal to the external price ratio. The country then exports PK of x and imports KT of y, thereby becoming better off than in the pre-trade situation, as the reaching of a higher indifference curve illustrates.[1]

It can also be demonstrated that the equilibrium situation under free trade fulfils the marginal, stability, and total conditions of the new welfare economics.[2] We thus indicate the gains from trade without any reference to real costs, since we concluded it was best to try and avoid their consideration if possible.

Nevertheless, as Shove remarks, 'You may drive out real cost with a pitchfork, but it has an awkward way of coming back.' And so it is here. To abstract from real costs it is necessary to assume that the factors are indifferent as to whether they are employed in producing x or y. Otherwise a change in production as a result of the opening of trade may involve an increase in real costs. One way of meeting this difficulty is to go back to the derivation of the p.p.c. We must then make allowance in the box diagram for any aversion which a factor may have for a particular occupation. Thus if labour should be averse to working in industry y, we should consider how much of commodity y would have to be given to labour in order to make it indifferent to working in x or y. This will result in a p.p.c. which, if it recognizes such compensation, will be below and to the left of the p.p.c. based solely upon the maximum output principle. Such compensation of factors involves nothing more than the general point that compensation will have some influence on production and that a complete analysis must consider the total effect.

---

[1] The cases of decreasing costs and constant costs could also be analysed in a similar manner, though in these cases specialization would be complete. The analysis can also be generalized to include more than two commodities and more than two countries, but no elucidation of the fundamental principle is gained thereby, and such complications need not be introduced here.

[2] Cf. J. R. Hicks, 'Foundations of Welfare Economics', *Economic Journal*, xlix (1939), pp. 704–5; A. P. Lerner, 'Cost Conditions in International Trade', *Economica*, xii (1932), p. 356.

The foregoing type of analysis is instructive in many ways. A symmetry of method is maintained when marginal, stability, and total conditions are applied to an open economy as well as to a closed economy. Pleasing as this symmetry may be, nevertheless it is questionable whether the welfare problems which are most significant for an open economy are sufficiently accented by such an analysis. When we merely expand our model to include the entire world economy as our relevant economic system and then apply marginal, stability, and total conditions to each sub-economy of the system, we are likely to suppress welfare problems which are peculiar to international trade.

This is especially true when we consider the welfare of one country instead of the entire international economy, since it has long been recognized that free trade need not be the best form of trade for maximizing the economic welfare of one country considered alone.

Yet if we concentrate attention only on the marginal and stability conditions we may overlook completely the welfare implication of a tariff. For, as Scitovsky shows,[1] a tariff leaves the allocating efficiency of single countries unaffected. The marginal conditions indicate that allocating efficiency for a country is at its optimum when the relative prices of all commodities are the same for all members of the community. A duty on international trade, though it would make relative prices different for the members of different countries, and hence reduce the allocating efficiency for all countries taken together, need not affect the efficiency of each country taken separately. For the prices relevant for any person always include his own country's tariffs and exclude those of all other countries.

Rather than multiply individual examples of this character, we may examine four general limitations of the preceding analysis:

1. It is frequently deceiving to draw analogies between a country's and

---

[1] T. de Scitovsky, 'A Reconsideration of the Theory of Tariffs', *Review of Economic Studies*, ix, No. 2 (1942), p. 93.

an individual's economic behaviour. This is especially true for considerations of demand, as a comparison of community and individual indifference curves will indicate. Non-intersection is a geometric property of an individual's indifference curves. But this need not be true for community indifference curves unless the economy is unified to such an extent that the indifference curves can be considered as those of the 'planners' or unless each member of the community has identical tastes and the same resources. Otherwise through a point, say, T in Diagram II, there is not one but an infinite number of community indifference curves, corresponding to different distributions of income among members of the community.[1] And since the community indifference curves need not represent the same distribution of income they may actually intersect one another. In the preceding analysis it was tacitly assumed that the curves do not intersect. Once this assumption is removed, the analysis is severely weakened, since it then becomes impossible to state that of two situations one is better than the other if it lies on a higher indifference curve, for each situation will be worse (or better) than the other for its own income distribution.

Nor does Kaldor's method of treating the community indifference curve as representing constant real income at every point along the curve eliminate this problem. For the index number problem involved in this method is identical with the possibility of intersecting community indifference curves.[2]

Furthermore, though the compensation principle has definite meaning in terms of a central government within a given country, when applied among countries there is the simple, but puzzling, question of locating the source which is to pay the compensation.

Finally, the powers of legislation, taxation, &c., open up a new province in which a country can influence trade in a different way and to a greater degree than can an individual.

2. The marginal, stability, and total conditions are also limited in their welfare connotation if they are taken to be simply the equivalent of a purely competitive solution. It is true that these conditions are fulfilled when pure competition exists. But various writers have demonstrated that the equilibrium conditions of pure competition need not be identical with welfare conditions.

Even more fundamental is the question of whether the purely competitive solution could be realized for the international economy. Whereas it is possible to argue that for a single country a certain type of central planning could institute the basic conditions and results of pure competition, the lack of an equivalent central planning body for the international economy makes such reasoning inapplicable. Even though each sub-economy achieved the purely competitive solution, it would not follow that a purely competitive solution would have been achieved *pari passu* for all countries together. The existence of spatial and natural differences among countries, with resulting specialization, would continue to make products of various countries not perfectly interchangeable—a condition contrary to the assumption of pure competition. In the 'real world' the existence of selling costs and the demonstrated preferences of consumers between imports and imperfect home substitutes show the existence of monopolistic competition in the international economy. Is there any reason to believe that pure competition within each country would *ipso facto* make the goods of various countries perfect substitutes?

---

[1] Cf. ibid., p. 94; and Viner, op. cit., p. 523.

[2] T. de Scitovsky, 'A Note on Welfare Propositions in Economics', *Review of Economic Studies*, IX (1941), p. 87.

In the absence of perfect substitutes, a country's reciprocal demand curve will not be perfectly elastic (even if we assume that the price-elasticity of the other country's supply of the first country's imports approaches infinity). For the other country the optimum tariff would then have some value larger than zero, thereby destroying pure competition between the two countries.

3. The foregoing analysis is also limited in so far as it restricts itself to comparative statics. Important welfare considerations may be involved in the changes which occur along the path taken to the new equilibrium position. To compare discrete equilibrium situations is not completely enlightening when it is observed that gains and losses occur over time and that these intertemporal lags and changes may have significant welfare implications.

4. Finally, marginal analysis in itself may have limited relevancy for welfare problems. To the basic question of whether a country should trade or not it has no answer. For this is a matter of total conditions. Indeed, Ricardo was concerned with total conditions in his discussion of which commodity should be exported and which imported. The marginal conditions, how far trade should be carried, first appear in significant form in J. S. Mill's exposition of reciprocal demand. In contrast to the Smithian and Ricardian interest with the 'widening' of an economy through increased physical productivity and expansion of the market, the marginal conditions relate to the 'tightening up' of the economy and become criteria for testing the allocative efficiency of given resources.

However, even overlooking the possibility that the fundamental assumption of continuous functions may be mistaken, it is still true that the marginal conditions, implying as they do a logic of infinitesimals, do not justify a like treatment of large and small changes in an economic variable. To divide a large change into small parts for the purpose of marginal analysis may involve a complete disregard of the fact that the change may be so great that it alters the entire system in a qualitative manner.

Moreover, it is too much to hope that the fine balancing involved in marginal analysis can or will be done by governments. Important as such balancing is in expressing the nature of the problem, the logic of infinitesimals is not likely to be the type of logic which pervades administrative machinery. The size of gains and losses with which actual policy is likely to be concerned will be of a much larger magnitude than marginal magnitudes.[1]

Thus, when it is a question of applying welfare economics to policy, the foregoing limitations of general optimum analysis give credence to the view that it may be important to be unimportant. Instead of concentrating on general welfare analysis with a fulfilment of marginal, stability, and total conditions, we may be able to proceed farther with international trade problems by using a case method involving partial analysis. The partial analysis will introduce the Marshallian concept of consumer's and producer's surpluses in their rehabilitated forms.[2] This concept is important not only because partial welfare analysis is grounded upon it, but also because, to avoid interpersonal comparisons of utility, we must employ the compensation principle of the new welfare economics which is but an extension of the surplus analysis.

4

International trade literature has long recognized some aspects of a partial welfare analysis. Marshall, for example, observed that in many cases a partial, and not a general, equilibrium approach might be more suitable. Referring to Ricardo's treatment of general taxes on the trade between two countries, Marshall observes that 'evils . . . may arise from a crude application of its suggestions to real problems. For that simplicity which makes it helpful, makes it also deficient and even treacherous.'[3]

Marshall, however, gave little attention to the partial approach. Instead he concentrated on the total 'net benefit' from international trade, since the concept of net benefit was in harmony with his reciprocal demand analysis. Following a method analogous to that of measuring consumer's surplus, Marshall determined from the 'trading schedules' that the aggregate benefit of a country's entire foreign trade consisted of so many bales or 'unit products of her labour and capital'.[4]

Edgeworth also saw the aspect of surplus in trade, but he too sought a total utility measurement. [1]Schüller, Barone, and Auspitz and Lieben, however, clearly recognized the partial analysis and its surplus implications. Yet, as Viner indicates in his *Studies*, their analyses are subject to criticism. But with the rehabilitated forms of consumer's surplus I believe we can avoid these criticisms while reinterpreting and extending these earlier analyses. Following the technique Marshall used in his general value theory instead of that in his international trade theory, we may consider only particular surpluses arising from trade in particular commodities. It now remains to indicate this alternative approach in greater detail.

---

[1] Cf. Myint, op. cit., p. 194.

[2] Hicks, 'Rehabilitation of Consumer's Surplus', *Review of Economic Studies*, viii, No. 2 (1941), pp. 108–16; id., 'The Four Consumer's Surpluses', ibid. xi, No. 1 (1943), pp. 31–41.

[3] Alfred Marshall, *Money Credit and Commerce* (1923), p. 190.

[4] Ibid., pp. 161–3.

The fundamental question involved in examining the relation between international trade and economic welfare is whether free trade is the optimal form of trade from (*a*) a cosmopolitan view, (*b*) a national view.

We have seen that the Ricardian comparative cost analysis answers this completely in the affirmative. But the Ricardian argument is weak on both the production and demand sides. We have also noticed how the modern indifference-curve analysis attempts to remove these debilities. Though it is successful in abstracting from a labour theory of value on the production side, nevertheless it can offer but little amplification of Ricardo's views on the demand side, due to the limited validity of community indifference curves.

It might be maintained that we do not need community indifference curves, but only knowledge of the equilibrium condition that the value of imports must equal the value of exports. It is quite true that this equilibrium condition will show the amounts imported and exported, but it will not allow a welfare conclusion unless the price line through this equilibrium point is tangential to a community indifference curve which is higher than the pre-trade curve. For it is possible that trade will involve not more of every commodity in comparison to the pre-trade position but rather more of one commodity and less of another.

Does this mean that no answer can be given to our initial question? I do not believe we need be forced to such an ineffectual position. For with appropriate modifications of some aspects of the preceding analysis we shall be in a position where we might apply a suitable partial welfare analysis to problems in international trade.

Let us examine a model consisting of two countries, trade in two commodities, and constant returns to scale in each country. This model emphasizes the Ricardian case of complete specialization—a case which is still extremely relevant in the modern international economy. For the existence of monopolistic competition with quality differences in commodities makes complete specialization common. Moreover, even though each country may be able to produce the same commodity under increasing costs, the absolute range of costs may be so wide that complete specialization will be undertaken by one country. Finally, as already noted, the less specific any one factor is, the flatter will be the p.p.c. Since multiplicity of factors will reduce the specificity of any one factor, it is not unreasonable to consider a linear p.p.c. when we recognize multiple factor production.

---

[1] F. Y. Edgeworth, 'On the Application of Mathematics to Political Economy', *Journal of the Royal Statistical Society*, lii (1889), pp. 553–60.

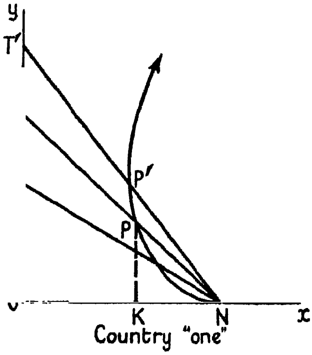

DIAGRAM III

Thus we begin with a linear p.p.c. for each country. By considering the distances from the origin at which the p.p.c. (*NM* in Diagram III) cuts each commodity axis we may determine in which commodity the country will specialize. Then the external price ratio or terms-of-trade line (*NT*) may be drawn from the relevant end of each country's p.p.c. Just how far a country will trade along this line will depend upon the workings of its internal price system. But we know that there must be some point *P* along the line *NT* which the forces of supply and demand will determine to be an equilibrium point in the sense that the value of imports equals the value of exports.[1] We are hence able to determine the amount exported (*NK*) and the amount imported (*KP*) of each commodity for each country—at given terms of trade (*NT*).

If the terms-of-trade line changes to *NT'* a new equilibrium point *P'* will be reached. Consequently, by varying the terms of trade and noting the corresponding amounts of commodities exported and imported we are able to derive an offer curve (*NPP'*) for each country, the offer curve simply being the locus traced out by the equilibrium point as the terms of trade vary.[2] These curves, while having the general shape of the Marshallian reciprocal demand curves, will not be identical with them, since they do not represent a similar general equilibrium type of analysis. But Marshall himself indicated *en passant* that

> 'by giving a special interpretation to the curves we may cause them to correspond closely to the actual circumstances of some important practical problems. . . . Instead of using distances along *OX* to measure cloth as representative of the whole of *E*'s exports to *G*, let us use these distances to measure only one of the wares which *E* exports, viz. coal. *OE* can no longer be called *E*'s demand curve but may be called *E*'s coal export curve; and *OG* may now be called *G*'s demand curve for coal.'[3]

---

[1] Samuelson, *Economics* (1948), p. 547.

[2] Cf. W. W. Leontief, 'Use of Indifference Curves in Analysis of Foreign Trade', *Quarterly Journal of Economics*, xlvii (1933), fig. 2, p. 494. Our procedure is similar to Leontief's, except that no use is made of community indifference curves.

[3] Marshall, *Money Credit and Commerce*, p. 357.

When we adopt this procedure and no longer refer to bales, but to actual individual commodities which are traded, we can proceed to derive demand and supply curves of the conventional shape. But to derive price–quantity diagrams we must divide the ordinate of *G*'s demand curve by the quantity of coal, since the ordinates along the reciprocal demand curve represent total receipts and hence correspond to areas under ordinary demand curves. Once we have ordinary demand–supply curves, we can proceed to apply a partial welfare analysis to determine if free trade in these particular commodities is the best form of trade for all countries together and for a single country taken alone.

It has frequently been demonstrated that from the cosmopolitan viewpoint free trade will create a situation in which it would be impossible to increase one country's welfare without diminishing another country's welfare.[1] But from the viewpoint of a single country it can also be shown that, under certain conditions,[2] a tariff on a particular commodity may increase the tariff-imposing country's welfare in the sense that the sum of consumer's surplus, producer's surplus, and government revenue might be greater with a tariff than under free trade. Though some individuals within the country may be worse off as a result of the tariff, nevertheless it may be true that others are so much better off that they might compensate those who are worse off and still be better off than they were in the pre-tariff situation. Similarly, those who will be worse off in the post-tariff situation could not bribe those who will be better off to a sufficient degree to make them as well off as they would be in the post-tariff situation.

If the tariff be a prohibitive one, so that none of the commodity is available, the loss involved is measured by the compensation which has to be given to consumers to make up for their loss of opportunity to consume the missing product, plus the compensation which has to be given to producers to make up for the excess of their earnings in the industry whose product was previously exported to pay for the import, but now contracts, over what they could earn in other uses.

The tariff example is illustrative of the character of the analysis which is applicable. We need not consider all hypothetical welfare distributions but merely those which exist before and after the contemplated change. Since the important issue is the movement from one situation to another as a result of a change in a particular economic variable, the rest of the system being assumed to remain constant, the significant criterion is that of the relative surpluses.

---

[1] Cf. T. de Scitovsky, 'A Reconsideration of the Theory of Tariffs', *Review of Economic Studies*, ix, No. 2 (1942), pp. 89–90.

[2] Cf. Stephen Enke, 'The Monopsony Case for Tariffs', *Quarterly Journal of Economics*, lviii (1944), pp. 229–45, for one particular case when the importing country acts as a monopsonist and discriminates against foreign producers. It should be recognized that Enke has ably demonstrated the use of partial welfare analysis in the monopsony situation. The present paper merely attempts to place this type of analysis in the larger context of comparative costs and to suggest extensions of the method.

We need not enter here into an analysis of the shades of difference between the Marshallian surplus and the four consumer's surpluses developed by Hicks. Each of these may be the more appropriate concept in certain situations. It is worth noticing, however, that since what Hicks calls the quantity variation is important in problems involving the disappearance or introduction of a commodity, the quantity variation may be the significant criterion in questions of a prohibitive tariff, discriminating monopoly of Pigou's first degree, and for analyses involving potential exports or imports. On the other hand, what Hicks calls the price variation may be relevant for problems involving a revenue tariff or certain types of quotas. But since Marshall's measure is intermediate between Hicks's four measures and coincides with them when their differences are negligible, Marshall's measure is a sufficient approximation for most purposes and the Marshallian demand curve may be used.[1]

And so we could go on showing how various situations in international trade which raise the question of economic welfare can be interpreted by means of the surplus technique and a partial welfare analysis. By this method it may be possible for the economist to say much regarding the welfare effects of tariffs—whether the tariff be for the purpose of taking advantage of decreasing costs, to achieve the gains of a discriminating monopolist, or those of a monopsonist. It should also prove meaningful for problems of dumping, barter agreements, quotas, and export subsidies. But detailed analysis of these problems is properly part of another story.

The moral of the present story is simply that analytical techniques are available by which we can proceed beyond the limited classical conclusions regarding the gains from trade. We may do so, not by means of the indifference approach and the general optimum conditions which we have seen also to be deficient in several respects, but by means of the more relevant partial welfare analysis. The value of the latter method is not only the pragmatic one of being more suitable for policy applications. It may also be quite pleasing to the economic theorist by virtue of its capability of encompassing several doctrines which have remained isolated from each other in the historical body of economic literature. Thus the method suggested in this paper incorporates Ricardo's important emphasis on what amounts to total conditions. It brings the doctrine of comparative costs, upon which the classical analysis of the gains from trade rests, into harmony with the general theory of value. Finally, the rehabilitated form of consumer's surplus now allows us to reinterpret in an unobjectionable manner the problems raised by earlier writers—as well as providing a basis for proceeding to an analysis of international trade problems which these writers overlooked.

---

[1] J. R. Hicks, 'The Generalised Theory of Consumer's Surplus', *Review of Economic Studies*, xiii (2), No. 34 (1945), p. 68.

UNIVERSITY COLLEGE,
OXFORD.

OLIGOPOLY AND IMPERFECT COMPETITION

By E. G. DOWDELL

I. Causes of Imperfection of Competition

IMPERFECTION of competition is generally due to three factors. In the first place, mobility is limited, so that no consumer can regard a large number of grocers' shops as equally convenient for the purchase of sugar, while the shops have unequal costs if they deliver to their customers' homes. Secondly, there are variations of quality, even in a commodity such as sugar, but far more in the case of most goods, especially if the conditions of sale and other ancillary elements are taken into account; so that among the most accessible sources of supply there are many causes why consumers should not be indifferent. In the third place, it is essential that, if they are to reach a level of efficiency that is at all tolerable, firms should be of not less than a certain size in respect of any one activity;[1] so that the number of suppliers cannot be indefinitely large.

Imperfect competition must usually be explained by a combination of these causes.[2] This can best be seen by supposing each of them in turn to be absent. Were people infinitely mobile, everyone might have access to a large number of restaurants of a particular type. If all biscuits were identical, there might be plenty of sources of supply in quite a limited area. Could firms be indefinitely small, there might be a large number in any neighbourhood offering each kind of biscuit. Inertia, causing 'goodwill', often reinforces the first two causes, but, taken by itself, is of little significance.

II. The Concept of an Industry

The great pioneers in the field of imperfect competition,[3] though they introduced many highly important types of analysis and pursued them with outstanding skill, did not emancipate themselves from all the special assumptions connected with the theory of perfect competition. This is not surprising; but it is perhaps permissible to use that epithet with regard to the extent of their continued reliance on the concept of the 'industry' or

---

[1] It is therefore not an answer to say, for example, that shops can combine the sale of a very large number of different goods, so that the need for shops to be of appreciable size does not prevent their operations in any one line from being small enough to permit the co-existence of a large number of competitors.

[2] If words are given their customary meanings, the qualifying adverb is necessary, as there are some exceptions. For instance, the total demand for lighthouses and crowns is so limited that, even if they were homogeneous, there would not be scope in the whole world for a large number of efficient producers. However, the term 'commodities' could without very great distortion be used in a sense which would justify a categorical statement.

[3] Robinson, J., *The Economics of Imperfect Competition*; Chamberlin, E., *Monopolistic Competition*.

'group' as employed by the predecessors from whom they were trying to break away. This led them down many wrong turnings and caused a great waste of ingenuity. Although there has been much criticism of this point of view,[1] it still seems to occupy a quite important place in the thought of most writers on the subject.[2]

In the analysis of perfect competition the 'industry', with its correlative, the 'market', connotes a group of firms each of which competes equally with all the others. The effect of a change in price and sales by any one is therefore diffused and makes an infinitesimal difference to each of the others. The boundaries of the group are quite definite, since a firm whose competition is in any degree less is to be excluded. The question how many firms there are in the industry is capable of a definite and significant answer.

Where competition is obstructed, the situation is very different. The first two main causes of imperfection create a chain relationship, or rather a complex of intersecting chains. Distance is not equal between one public-house and all others in a 'group'. It therefore may damp only slightly the competition between each and two or three others, while placing many quite beyond reach of direct competition. Yet those many cannot be simply ignored, since some of them compete appreciably with houses that affect the one under consideration. If one kind of biscuit is rationally preferred to all others, how can the degree of preference be uniform, when the differences in sweetness and other flavours, size, crispness, and so on are obviously unequal? Two sorts of chocolate biscuit are generally fairly good substitutes for each other, but a big gain on cheapness (or points value) will induce many buyers to shift to numerous other flavourings, and we have no clear guidance as to where this tendency becomes negligible. It certainly is often significant in regard to things which are not called 'biscuits', such as cakes and chocolates. The exponents of the traditional theory of imperfect competition admit this difficulty with more or less candour, but rely on a 'gap in the chain of substitutes'[3] which will show us where to stop and give a definite grouping. Yet it may be said, almost without qualification, that gaps of sufficient width to be serviceable for the purpose are imaginary. In considering the mobility of sugar consumers over distances of one, ten, a hundred, or a thousand miles, on foot, by bus, private car, train, ship, or aeroplane, where is the obvious *terminus ad quem*? If cakes, cigarettes, even bread may sometimes be regarded as substitutes for a particular brand of chocolate biscuit, how are we to define the 'commodity' in something better than a merely verbal fashion?

---

[1] Especially by Triffin in his *Monopolistic Competition and Equilibrium Theory*.

[2] Chamberlin (op. cit. v. 3) admitted that the imperfection of competition is not uniform throughout the market corresponding to the group under consideration. The differentiation of the product is not uniformly spaced. If high average profits attract new competitors, very different effects will be produced on different members of the group. Yet he thinks that these variations will produce no real difficulty in the end. He therefore proceeds under the 'heroic' assumption of uniform demand and cost curves for all members of the group. He also imagines, consistent with differences between the products of all sellers, possibilities of product variation which are uniform for all, though his concrete instances adduced to justify this procedure do not seem at all convincing.

[3] Robinson, E. A. G., *Monopoly*, p. 6. Chamberlin (op. cit., pp. 103–4) notes the chain relationship, but is still inclined to look for definite groups and sub-groups.

Unless we have to do with complete monopoly,[1] we find that the factors preventing perfect competition erect barriers of very unequal strength. There may be a very high degree of competition between one firm and two or three others, a smaller degree with a few more, and so on. There is a continuous gradation as we go farther from our starting-point in geographical distance and in qualitative change. Every firm has 'neighbours'[2] who may be really close, and it is the essence of this relationship that it should not cease abruptly at a particular line. A firm which cuts its price, improves its product, or advertises more vigorously will attract much more custom from some producers than from others. If the amount is large by absolute as well as by relative standards, then the neighbourhood factor is of great economic significance and creates an oligopolistic relationship. If one firm can increase its sales substantially, a large part of the increase will be by diversion from rivals, not growth of consumption, and the diversion from one or two rivals will be great enough to raise the question of their reaction in self-defence. They must be expected to lower their price, improve quality, or advertise more; so that the firm initially contemplating an attempt to sell more cannot take a rational decision without forming an opinion as to the probable nature of this reaction. The supposed marginal revenue depends upon this estimate, which may be wide of the mark, as well as upon the view taken as to the response of consumption.

III. Perfect Oligopoly

It may be hazarded that the wish to avoid the problems of oligopoly has to a great extent been father to the thought that usually it is unimportant.

---

[1] It may be questioned whether there are cases of complete monopolies. Perhaps an intermediate case is of some importance. Where there is a substantial gap, but one not large enough to make competition negligible, there may be half a dozen or more firms close to the other side of it, various factors combining to make their products about equally substitutable for those of the firm under consideration. The oligopolistic relationship may then be considered unimportant; but, if it is, there is another important consideration that is usually neglected. When a firm can attract the least loyal buyers from a considerable number of rivals, its demand curve will look decidedly elastic below the ruling price, so that, when price is much above marginal cost, there will be an incentive to lower it even though the amount of trade taken from any one other firm is small enough to suit the traditional type of analysis. On the other hand, a price increase will deprive the firm of a much larger fraction of its trade than a similar decrease would draw from any one rival. In such circumstances, the usual marginal revenue curve is inappropriate. Cf. Nicols, 'The Rehabilitation of Pure Competition', *Q.J.E.* lxii, pp. 31 et seq.

[2] The word is Mr. Kaldor's—'Market Imperfection', *Economica*, 1935, pp. 39–40.

This is not a matter for surprise when we consider how much the analysis depends upon what view is taken of the reaction of competitors and how little justification there is for the assumptions that have been employed in order to reach definite results.[1] The idea that competitors will maintain their outputs unchanged in face of a price-cut by one firm seems to have nothing to recommend it, save in the fanciful case of adjacent mineral springs whose water can be produced and marketed without cost. Opinions may differ as to the precise relation between cost and price, but the implication that they have no connexion whatever is certainly not acceptable. The notion that competitors keep their prices stable seems more plausible than that they reduce them just enough to maintain sales, because it has the force of inertia in its favour in the many cases where the producers' deliberate decisions are about price rather than output. But this is only a short-period consideration. When sales alter owing to a competitor's price change, it is only reasonable to adhere to the old price if costs, marginal or average, are unaffected by the amount produced, which is, to say the least, a rare case. Indeed, where oligopoly is perfect, that is where the cross-elasticity of demand is infinite, the only plausible assumption is that prices of all the competitors are the same, so that a change by one must be exactly followed by the others. Otherwise, either one or more will be driven out of the market completely, or those charging less will have to limit their sales by some means other than prices and will consequently lose opportunities of profiting by a higher price unaccompanied by reduced turnover. This does not mean that we have a determinate solution, unless a price is established which all happen to regard as the most desirable. This will not occur if they have different costs or different notions as to the elasticity of total demand.[2] There is consequently likely to be complete instability as each in turn initiates a price change and is followed by the others. This result is very unsatisfactory for the theorist.[3]

If, in spite of the considerations already advanced, it is still thought that firms envisage the possibility of other policies on the part of rivals whose business is affected by a price change, then we must allow for uncertainty instead of—or as well as—instability. A firm may then decide upon a change in the light of a wrong estimate of the reaction it will provoke, other firms may decide their responses under a misapprehension of the first firm's further reply, and so on.

---

[1] It is convenient to take duopoly as a special case of oligopoly. We can then regard Mr. Kahn's article on Duopoly (*E.J.* 1937, pp. 1 et seq.) as one of the best discussions of the problem. Prof. Chamberlin (op. cit., ch. v) also deals very skilfully with many of the difficulties.

[2] It is reasonable to suppose that, as the market is shared in a random way, each will tend to have the same fraction. If their cost functions are dissimilar, it is conceivable that customers might happen to be distributed unequally in such a way as to make marginal costs all equal, but this would be a very rare case. It should perhaps not be stipulated that all sellers take the same view of elasticity of demand, as it will be sufficient if they do not hold conflicting views strongly.

[3] Solutions through price leadership (as in Kahn, loc. cit.) require rather special assumptions. It is not clear how Mr. Kahn would meet the objection that, if the follower's price is far enough below the leader's to make all buyers prefer him, he is missing an opportunity of increasing his profit by raising the price without reducing demand below the amount whose marginal cost equals that price. Actually, it is usual to think of the leader as the larger and stronger concern, in which case the follower can expect to sell as much as he desires without charging less, being content with less than half total sales.

It is probable that under such conditions there would in practice be a strong tendency to price rigidity. Where the response of competitors is not known and may be of a kind leading to what can properly be called a price war, adherence to a ruling price has much to recommend it, if it yields fairly satisfactory profits, especially for business-men who rank security high among economic desiderata and are glad to avoid the trouble of constant price changes. Mr. Kahn characterizes this as 'the first degree of collusion' on the ground that one firm cannot be sure that others will refrain from price-cutting if it does so.[1] This is a dubious argument, however, seeing that, in a perfect market, the advantage of being first with a reduction is negligible, as consumers lost will be easy to recapture. Waiting and seeing is therefore not dangerous. In an imperfect market, on the other hand, the loss of trade immediately following a rival's reduction will be limited and, if a business-man fears that he might not in the ordinary way hear of a cut in time to prevent serious damage, it must usually be more sensible to arrange for early warning than to lower his own price as a precaution. Oligopoly, even without collusion, is therefore almost certainly an important cause of the increasing rigidity of prices. This is naturally not to deny that there is a strong motive for concerted action to prevent changes that would occur in its absence. Without collusion, a 'monopoly price' is possible, though not probable, as there is also uncertainty about the response to price increases, following the establishment of any price lower than a monopolist would choose. That deters anyone from starting an upward movement.

IV. Imperfect Oligopoly

The difficulty, however, is in no small measure created by the theorists. Perfect oligopoly has probably never existed, for it is extremely difficult to conceive of any factor or combination of factors which could divide a small number of firms sharply from all others, yet leave them in unrestrained competition with each other. It should be emphasized that a pronounced gap is required, for otherwise other firms, perhaps a large number of them, will affect and be affected by those under consideration, whereas it is characteristic of oligopoly analysis that the effects of price change are highly concentrated. If there are four tea merchants so close together that the geographical barriers may be ignored, but others also fairly easy of access by personal call, post, or telephone, a price increase by one may not be copied by the others for fear of losing custom to outsiders. For this reason, as well as because he will himself lose trade to the outsiders, any one seller would hesitate to initiate a rise. A ruling price appreciably above marginal cost will be unstable because he will expect a small reduction to increase his sales greatly, diverting varying amounts from a number of sellers. In other words, he regards the demand for his output as highly elastic, so that something not far from the competitive solution is indicated. There will be little uncertainty or instability. The same argument would apply to a small number of manufacturers of a product of uniform quality, with others making not very dissimilar articles. It may be that instances could be found where there are wide enough gaps to make the group so self-contained that the oligopoly problem is found in virtual purity, but it is suggested that these are rare.

---

[1] Kahn, loc. cit., p. 9.

Much more often there are some barriers between a firm and even its closest competitors. Then if it lowers its price, not nearly all their customers will try to buy from it. Adherence to the old price then becomes a possibility for the other firms, their decision naturally depending on the amount of sales they will lose, which will be inversely proportional to the importance of the barriers. Moreover, if the old price no longer seems the best, it does not follow that an identical cut must be made. A smaller reduction may be sufficient to reduce the diversion to small dimensions. If the total change is divided into a number of smaller changes, the advisability of reductions of various amounts could be thought of in terms of the loss of receipts from sales which would be made even without a small cut, balanced against the advantage of retaining business which would otherwise be lost. Normally it will not pay the competitors to go the whole way down. Similarly, if one firm raises its price, only some of its customers will transfer, so that the other firms will not be overwhelmed with orders and must seriously consider the advisability of maintaining the old price and doing more trade. If they decide on an increase, it will generally not be by the full amount. So far it may be objected that, if the instability of the market as a whole is reduced, the element of uncertainty is increased; this is supposing that, under perfectly oligopolistic conditions, each would confidently expect the others to copy any move he made. If other solutions are still considered reasonable, then there is uncertainty in any event. However, it must be observed that the few firms competing imperfectly with each other will usually also compete appreciably, though less keenly, with others. A price rise will cause some diversion to them, but not enough to make any one of them likely to follow in an upward direction. Thus the firm being studied could not rationally consider so large a rise as if its effect were more concentrated. Likewise, a reduction will win customers from many firms, so that, unless it can increase sales indefinitely at constant marginal cost, the size of the most profitable cut is lessened.

It might be thought that, if firms *B* and *C* compete with *D*, *E*, *F*, &c., as well as with *A*, the likelihood of their following *A* downward is increased. But their motive for diverting trade from these others is independent of an initial cut by *A*. There will be some critical degree of separation between *A* and *D* which will lead the latter to stand by his own price unless *B* or *C* also cuts by the same amount. A smaller cut by *B* will enable him to attract some of *D*'s customers and retain some of his own who would otherwise go to *A*. *C* may make a still smaller reduction. If *B* follows *A* exactly and the chain relationship is continuous, then *C*, *D*, *E*, and so on will do likewise. *A*, for his part, will under many conditions be deterred from making a larger reduction by the desire not to disturb *D*'s price, even though he regards the possibility of gaining some custom as a reason for making a cut which would not appear profitable if the only possible gain was from *B*. It seems, then, that the imperfection of competition between even the closest rivals, coupled with some competition with a considerable number of others, has something of an anchoring effect, reducing either instability or uncertainty or both.[1] It may destroy the complete temporary stability among nervous oligopolists, by lessening the fear of a really violent price war. In fact, it may be said to make small price reductions easier and more probable for the very reason that large ones are not likely. The reaction between the closest neighbours is damped, still more its spread farther afield. The smaller the defensive reduction which one firm expects others to make, the less is the possible error in estimating its size, so that there is less room for indeterminateness[2] or, if that is preferred, determination by wrong judgement. This goes far to remove the possibility that a price change made by one firm because of an incorrect appraisal of the oligopolistic relationship will so disrupt the *status quo* that the behaviour of that market will for ever after be substantially different from what it would have been if this mistake had not been made. The damage done to equilibrium theory based on objective factors, as distinct from estimates, is correspondingly reduced.

V. Product Variation and Selling Cost

Oligopolistic interdependence has so far been viewed in relation to price-cutting. Product variation must also be considered. If one car is improved in a particular way, will makers of other sorts be led by the resulting fall of their sales to follow suit? Here the concept of 'perfection' is blurred, since, while all buyers may be presumed to prefer lower prices, it would be altogether unreasonable to assume that they all have the same attitude to the various possible improvements. Consequently, even if we started with a homogeneous commodity and something like a perfect market, it is extremely unlikely that all firms would make the same changes. They may fear the loss of much trade if they do nothing when one introduces an improvement, and they may therefore do something of the same order of importance, but allow him to attract the people with the strongest preference for his particular innovation, retaining most of their former customers and gaining some new ones by introducing an improvement differing in greater or less degree from his. Indeed, having been forced into action, they may seek to turn the tables on him by offering a more attractive bait. As before, there will be great uncertainty in action and reaction. Extreme instability is unlikely, however, as the abandonment of genuine improvements is much less common than the reversal of price concessions.[1] Starting from a highly imperfect market with barriers between all firms increases the number of questions upon which there is doubt, but greatly lessens the degree of uncertainty about each of the possible responses of competitors, because the latter are so much less disturbed. It must be borne in mind that, when a producer alters his product so radically as to become mainly a competitor of different firms, he must not only look to these for defensive measures, but must also consider the possibility that his old competitors, finding competition reduced, may take the opportunity to raise their prices, so that the loss of old customers by the innovator is diminished.

---

[1] Contrast Prof. Boulding's view as to the instability of imperfect oligopoly, *Economic Analysis*, p. 628.

[2] Cf. Edgeworth, *Papers relating to Political Economy*, vol. i, p. 121: 'It will readily be understood that the extent of indeterminateness diminishes with the diminution of the degree of correlation between the articles.'

Selling-costs also call for notice. Some of the most wasteful forms of advertising result from ill-judged efforts by one firm to expand sales without giving due regard to the likelihood that those firms which lose substantial amounts of trade will try to defend themselves. We often find heavy outlays on simple advertising and more complicated forms of sales pressure which appear to have no appreciable effect beyond maintaining the *status quo*. So far as profits are concerned, this may produce a 'competitive' outcome, in the same way as does price-cutting without regard to rivals' reactions, but in other respects, especially the all-important one of benefit to the consumer, the situation is very different, as it is a characteristic of heavy advertising to arise in imperfect markets and to make them still more imperfect. The consumers' ideas of what they are buying may be changed very much, even though their custom is shared in much the same way as it would be without advertising. Therefore we need not feel very much regret if producers come to an agreement to limit their advertising, abolish gift coupons, and so on. In itself, of course, this will not lower prices or yield goods of really better quality, but it will make some such genuine improvement more likely, and the direct loss to the consumer is usually small and may be negative. Informative advertising makes competition more perfect, but is discouraged by the oligopolistic relationship,[1] as the firm considering its advisability will have to allow for the possibility that, if a large number of previously ignorant customers are attracted from any one other firm, the latter will have to defend itself and may divert some old customers from the first firm, as well as lessening the latter's gains from others.

---

[1] Perhaps, therefore, we can neglect the case where a firm is considering a deterioration of its product and the effect of uncertainty as to how far rivals will deteriorate theirs when they find their markets increased as a result. This might be worth considering in connexion with some direct services. Owing to the diversity of tastes and possible improvements, it is improbable that the fear of reprisals ever completely prevents change, particularly as business men apparently regard quality improvements and advertising as less provocative than price-cutting.

VI. Entry and Profit

Among the consequences of the factors here discussed, one of the most important is that the concept of free entry, as usually understood, must be rejected.[2] Except where markets are perfect, it would be foolish for the would-be creators of new firms to regard the existence of excess profits as decisive. A new-comer must expect those most affected by his presence to cut prices, improve their products, or advertise more—or carry out some combination of these policies. His expectation of profit is therefore decidedly below what they have been earning, unless he has superior ability or other advantages.[3] It must be remembered, moreover, that temporary measures of a really 'cut-throat' or 'clubbing' variety are quite likely, since defensive action is most effective against a firm which has not yet had time to build up goodwill and financial reserves. There may be a good prospect of recouping their cost later, provided that they succeed. For a short time, therefore, threatened producers may lower their prices, improve their service, or advertise more than they would deem proper were the newcomer to gain a permanent footing. This clearly discourages efforts to enter, restricts the tendency to excess capacity, and facilitates the maintenance of excess profits. It is obviously more influential where production and selling cannot be effectively started on a small scale, so that important hostages must be given to fortune. By contrast, excess capacity is conspicuous in the retail sweet and tobacco trades. In addition to the possibility of starting business in them in a small way, the 'price-maintenance' policy of manufacturers limits the ferocity with which established shops can retaliate. In general the initial obstacle of heavy expenditure is, however, not as serious as it would be if new undertakings were not frequently established by old firms of great financial strength, built up, perhaps, in some quite different field.

---

[1] Strictly speaking, what is relevant is the degree of interdependence after, not before, the advertising. This is not an entirely unimportant distinction, as, for example, if a bookshop which had hitherto relied on local trade should contemplate entering a wider field by distributing catalogues all over the country.

[2] Cf. Triffin, op. cit. iii. 3.

[3] Prof. Chamberlin's frequent recognition of these problems did not save him from falling into grave error in this regard. He says, for instance (op. cit., p. 171), '. . . whenever price competition fails to function, *whether because each seller is in close competition with only a few others* or for any other reason, the result is not merely higher prices, but also excess capacity as a permanent and normal characteristic of the equilibrium adjustment' (my italics). Mrs. Robinson showed much less awareness of the problems involved. The argument advanced above that oligopoly makes for adherence to existing prices obviously does not apply when a fresh rival appears.

If we must abandon the expectation that surplus profits will be eliminated[1] by the entry of additional firms, we must as a result reject the argument that, even in imperfect markets, there is a strong tendency for price to equal average cost, including normal profits. This notion of the tangency of the average cost curve with the demand curve has had a great vogue, but, in so far as it has real meaning, it rests upon flimsy foundations. Normal profits, difficult enough to define with precision even under competitive assumptions, seem to elude the theorist when these are given up. Oligopoly, as well as other factors, tends to create a wide gap between the lowest which will keep existing firms in action[2] and the rate which will attract new-comers. Conditions are so diverse that there is little hope of generalizing, and the real issues have been obscured by desperate efforts to do so. Upper and lower limits to profit can no doubt often be conceived with more or less accuracy, and sometimes they may be close together; but it should not be expected that they can usually be given enough precision to be embodied in a curve purporting to show average cost accurately. In seeking to explain prices, we must throw what light we can upon the views of business-men regarding the prospects of establishing themselves in various fields and on the appraisals by existing firms of the attitudes of possible new rivals.[3] Nothing but harm, however, can result from the claim to reach exact conclusions.

ST. JOHN'S COLLEGE,
OXFORD.

---

[1] For the present purpose we need not say exactly how excess profits should be judged. All that is necessary is to rule out the tautological view that whatever profits are being earned are normal.

[2] A firm with experience of oligopolistic conditions will be reluctant to close down (among other reasons) owing to the anticipated difficulty of entering another line or even restarting in the old one.

[3] Cf. Andrews, 'A Reconsideration of the Theory of the Individual Business', *Oxford Economic Papers*, New Series, Vol. I, No. 1.

THE FOUNDATIONS OF WELFARE ECONOMICS

By I. M. D. LITTLE

PART I

The Criterion for an Increase in Welfare

THE formulation of the theory of economic welfare, which appears to be accepted by many economists to-day, consists of a number of elegant deductions from unacceptable premisses. This paradoxical state of affairs is the result of some logical and philosophical confusions. It is proposed, in this article, to discuss the criteria for an increase in the economic welfare of individuals and of society, and other related questions.

The criterion for an increase in the economic welfare of an individual over some earlier situation is that the goods of the earlier situation are in the later situation rejected in favour of those actually bought, i.e. the earlier set of goods could again have been bought, but was not. This criterion may be written for short $\sum p_2 q_2 \geqslant \sum p_2 q_1$. It seems that most people would probably be willing to accept this criterion. But, as thus stated, it is not adequate for the purposes of welfare theory. It must also be assumed that there are individuals who are perfectly consistent;[1] that is, roughly speaking, the pattern of their economic behaviour must remain unchanged so long as all prices remain unchanged. It is possible that eagerness to apply the theorems of welfare economics may have led economists to overestimate the consistency of individuals, and perhaps also to overestimate the extent to which the choice criterion of an increase in economic welfare is acceptable. But with these matters we shall not here concern ourselves. We shall assume that if 'on a higher behaviour line or indifference curve' can be taken to mean anything, then it is a good criterion for an increase in the economic welfare of the individual, so long as he is independent of the welfare levels of other individuals. That is all that welfare theory requires.

Welfare economics is primarily concerned with the welfare of the community, and not the individual. The latter is only important in so far as it is thought that changes in the welfare of a group must be defined in terms of changes in the welfare of individuals. Since 1939 writers on welfare economics have relied largely on the criterion for an increase in social welfare proposed by Mr. Kaldor.[2] This criterion is to the effect that social welfare has increased if a certain change has benefited at least one person and harmed no one, or, if some are harmed, then the beneficiaries must be able to compensate the losers without any of them themselves becoming losers, and at least one of them remaining a beneficiary.[1]

---

[1] I have sought to explain more exactly the meaning of consistency in a previous article: 'A Reformulation of the Theory of Consumer's Behaviour', *O.E.P.*, New Series, No. 1, Jan. 1949.

[2] N. Kaldor, 'Welfare Propositions in Economics and Interpersonal Comparisons of Utility', *E.J.*, Sept. 1939.

The salient feature of the new welfare economics is that it does not require the assumption that satisfactions are quantities which can, therefore, be added. In fact, different individuals' satisfactions are regarded as incommensurable, and the class, each member of which is the satisfaction of a different individual, is a heterogeneous collection. Now one such collection can only be said to be larger than another if some one item is larger and no item is smaller. This is the *only* meaning which can be attached to 'larger than' when one is dealing with incommensurable items. Therefore, if 'general welfare' refers to a collection of satisfactions, then one cannot, as a matter of logic, say that the general welfare has increased if anyone is harmed. Yet clearly there is nothing absurd about saying that the general welfare has increased when some people are harmed. It is frequently said, and there is nothing wrong with saying it.

From the above argument it follows either (1) that everyone who regards such a statement as 'the general welfare has increased' as meaningful is a utilitarian, in that he believes it makes sense to speak of adding different people's satisfaction, or (2) that 'the general welfare' does *not* refer precisely to a collection of satisfactions. Of these alternatives, the first is false, and therefore the second must be accepted. Also, since the criterion under consideration only provides *hypothetical* compensation for the losers, it follows that anyone who accepts it must either make interpersonal comparisons and decide that money has the same utility for everyone, or he must abandon the view that 'the happiness of the group' is a name for a collection of individual satisfactions. Of these alternatives the first again must be abandoned, because the new welfare economics prides itself on having got rid of the necessity of comparing different individuals' satisfactions.

Since, therefore, the happiness (or well-being) of the group is not the name of a collection of satisfactions, it must be otherwise defined; or rather, some instructions about the usage of the phrase 'increase in the happiness (or well-being) of the group' must be given. Now it may be claimed that this is just what the criterion under consideration does. An 'increase in general welfare' is to be regarded as a technical term, and the criterion instructs us how to use it.

---

[1] A. Bergson, in his important article 'A Reformulation of Certain Aspects of Welfare Economics' (*Quarterly Journal of Economics*, 1938), does not propose a sufficient criterion for an increase in economic welfare. He deduces the 'optimum' conditions from explicit value premisses as necessary conditions for a formal economic Utopia. On this formulation the only sufficient condition for an increase in economic welfare, which is implied, is that all the 'optimum' conditions, together with an 'optimum' real income distribution, be put into operation together. It is Utopia or nothing.

Since the new welfare economists are not utilitarians, the above position seems to be the only possible one which they can take up. Is there anything wrong with this position? The answer would be 'no' if the phrase under discussion was of the same type as, for instance, the word 'gravity' or the word 'force'. But it is not of the same type, and it is indefensible to claim that this is a perfectly objective scientific procedure. Such phrases as 'an increase in general welfare' and 'an increase in the happiness of society' and 'socially advantageous' are persuasive, influential, quasi-ethical terms. Any definition which one proposes for a phrase of this kind conceals an ethical judgement. There is no doubt that if one states that a certain change would increase the general welfare, or would benefit society, then one's statement implies that that change is *good*, and that it ought, other things being equal, to be made. By 'other things being equal' is meant 'assuming that no other ethical considerations are involved'. Thus it follows from the nature and usage of welfare language that welfare judgements are, partly at least, value judgements. It should be noted that this does not exclude the possibility that they may also be descriptive. One can, of course, both describe some change and imply that it is good. This is exactly what one is doing if one lays down a definition of 'an increase in welfare', and then uses the phrase in accordance with the definition. Thus, when a new welfare economist states that a certain change has increased welfare, he is stating that either someone has benefited and no one has lost; or that, if there are losers, the beneficiaries could compensate them, and at least one person end up as a beneficiary. What is not generally realized is that he is also saying that this is a good thing to have occurred. To put the matter shortly, one may say that any definition (or criterion) of an increase in the general welfare is bound to be a persuasive definition, or criterion, in just the same way as any definition of 'a gentleman' is a persuasive definition.[1]

There is, of course, no reason why we should not make persuasive definitions if we want to, but we should realize what we are doing. It has been claimed that the new welfare economics is objective and scientific, because it was thought that by rendering interpersonal comparisons unnecessary value judgements had been eliminated. As will be argued later, it is not only false that interpersonal comparisons are not required, but it is also false that interpersonal comparisons are primarily value judgements. They are quite certainly more scientific and more objective than judgements about 'the welfare of society'.

---

[1] Persuasive definitions were brought to light by Professor Stevenson, in an article of that title in *Mind*, March 1938.

We may now examine the question whether the criterion under consideration is acceptable or not. I think it could be regarded as acceptable if a large majority of people would be likely to accept it. (The mere fact that people might refuse to accept it is sufficient proof that it is a persuasive criterion.) Now it seems unlikely that many people would accept the proposed criterion, largely because it only makes compensation hypothetical. Thus, if a certain change makes the poor poorer, but the rich so much richer that they could compensate them, we must, by this criterion, say that the welfare of the community has increased, even if the poor are not, in fact, compensated. It is suggested that, in the prevailing climate of opinion, comparatively few people would be prepared to recommend such a change; yet to say that such a change would increase the welfare of society is to recommend it. The trouble is not that people would never say that welfare had increased when some people were harmed, but that they would want to know which individuals, or what class of individuals, was to suffer. The criterion proposed in no way identifies the individuals to be harmed. Everyone, except one person, might be harmed, and it would still be an increase in welfare.

Some objections may now be considered. The first objection is that the new welfare economics is about changes in *economic* welfare, not welfare in general. But to say that the economic welfare of society has increased can only mean that economic causes have increased the welfare of society. It may happen that some change, which would be called an economic change, would have both economic and political effects. These effects may be held respectively to cause an increase and a decrease in welfare—in which case one might say that economic welfare had increased and political welfare decreased. To say then that economic welfare has increased is, in effect, to say that the consumption of goods (in a very wide sense) and the work done have changed in a favourable manner. Thus, if a person argues that a certain change would increase economic welfare, it is open to anyone to argue that it would decrease spiritual or political welfare. Therefore, if the word 'economic' is inserted before the word 'welfare', the persuasive and influential effect of the statement is diminished. It is diminished, but not eliminated; after all, people still feel strongly even about economic welfare. Moreover, and this is important, while economists may begin by writing about changes in *economic* welfare—or, weaker still, changes in real income—the phraseology often gets changed after a few pages into the more forceful mode of expression. The frequent omission of the word 'economic' is not insignificant.[1]

[1] It should be noticed that, in the article cited, Mr. Kaldor did not explicitly propose the hypothetical compensation formula as a criterion for an 'increase in welfare'. He actually proposed it as a criterion for (*a*) 'an increase in physical productivity', (*b*) 'an increase in aggregate real income', and (*c*) 'an increase in the sum total of wealth'. Later, in reply to

The second objection is that even if what is here maintained is true, still a choice of colourless language can be made and, once again, the new 'welfare' economics would be respectable and scientific. This objection is valid, but it misses the fact that it is welfare that people are interested in. Some people, on the one hand, want to make recommendations, and others want to be told, without the responsibility of deciding for themselves, what is the best thing to do. If they are told without its becoming too obvious that they are being told, so much the better. It is, in fact, attractive to some people to have a theory which tells them what is the best thing to do. If the emotive and influential elements of welfare theory were removed, it would lose a great deal of its appeal, for people would then have to think how far the changes described really were relevant to their welfare judgements. A colourless transcription of 'this change would increase the welfare of society' would be: 'If this change were made, then we should have a situation in which it would theoretically be possible to put everyone (if everyone was like the perfectly constant and consistent 'individual' of economic theory) in such a position that they could choose at least what they had now. Practically, of course, it could not be done; some would gain and some would lose.' Most people would surely want to know more before they were prepared to say 'Yes, it would increase the welfare of society'. The theory of 'welfare economics' could certainly be stated in colourless language, just as Euclidian 'Geometry' need not employ spatial words. But, if the theory is to be applied, a normative interpretation is essential. Indeed, applying 'welfare economics' (that is, removing the inverted commas) is nothing other than giving it a normative interpretation. And if the theory is to be interesting in any other way than as a purely logical exercise, it must be interpreted. The choice is not between an uninterpreted abstract system, which employs colourless concepts so as not to suggest an interpretation, and a normative system, but rather between an acceptable and an unacceptable normative system.

We must now examine a defence of the Kaldor criterion which has been put forward by Mr. Hicks.[1] Mr. Hicks writes:

> 'In fact, there is . . . a perfectly objective test which enables us to discriminate between those reorganizations which improve productive efficiency and those which do not. If A is made so much better off by the change that he could compensate B

---

an article by W. J. Baumol, attacking the hypothetical compensation formula ('Community Indifference', *R.E.S.*, 1946–7), Mr. Kaldor wrote: 'The important point surely is that when the production of wealth goes up, some income distribution could be found which makes some people better off, and no one worse off than before.' This is not an important point, but an important definition of the phrase 'increase in the production of wealth'. All these four phrases are emotive and suggest, more or less strongly, that a change, such that the gainers by it could compensate the losers, would be a good thing. Why can't we just say that the change would benefit some people enough to compensate those who lost? Why must we go on to describe such a change as increasing the sum total of wealth, &c.?

[1] J. R. Hicks, 'The Rehabilitation of Consumer's Surplus', *R.E.S.*, Feb. 1941, p. 111.

> for his loss, and still have something left over, then the reorganization is an unequivocal improvement. . . . If the economic activities of a community were organized on the principle of making no alterations in the organization of production which were not improvements in this sense, and making all alterations that were improvements that it could possibly find, although we could not say that all the inhabitants of that community would be necessarily better off than they would have been if the community had been organized on some different principle, nevertheless, there would be a strong probability that almost all of them would be better off after the lapse of a sufficient length of time.'

It may be noticed that Mr. Hicks does seek to justify the criterion and that he is thereby, in effect, agreeing that one cannot simply define changes in welfare as one will. It may also be noted that he begins by giving a criterion for an improvement in productive efficiency, but ends up by saying that a reorganization which effects this is an *unequivocal improvement*.[1] The justification for saying this is that if all such 'improvements' were made, then, after a long time, almost everyone would be better off than he would have been otherwise. If this were true, it might be considered to be a good justification for saying that such a policy would, in the end, probably lead to an improved state of affairs, though not for saying that every change made under that policy was itself an improvement. But is it true? It is very dangerous to appeal to a long period of time, because over a long period tastes change, and the criterion for an individual's change in welfare becomes very shaky (or, if one insists too much on its rigorous interpretation in terms of indifference curves, it becomes meaningless). For the same reason, changes that would otherwise have been correct by the criterion may cease to be correct. Also, over a long period of time many people die, and, for this reason, the contention that almost everyone would be better off becomes rather absurd. Again, Mr. Hicks appeals to a whole series of such changes. But a whole series of the kind of changes envisaged would require political adjustments; so, when appealing to such a series of changes, it becomes important to insert the word 'economic' before the word 'improvement'. Finally, even if it could be said that there was a high probability that almost everyone would be economically better off in the end, it would not follow that everyone would agree that the result could be said to be an unequivocal improvement.

Nevertheless, there is clearly something in this justification. Whether it would be held to justify a series of changes or not would clearly depend on the magnitude of the expected potential 'gains'. It might, for instance, justify us in introducing important labour-saving inventions. Then, especially if the large losers were actually compensated, distributional effects might more or less cancel out. Admittedly, in some cases the compensation of large losers might be thought wrong in that they might be held to be the sort of people who ought to be made poorer. Prevailing ethical opinion would, on the other hand, probably think it better to make a redistribution of income in a less arbitrary manner.[1] Perhaps it might also be applied to a series of changes, such as making price equal to marginal cost in a large number of firms. It should be noted, however, that it only follows from 'welfare' theory that price ought to equal marginal cost if *all* other 'optimum' conditions of production and exchange are fulfilled. It is quite wrong to suppose that the theory implies that price ought to equal marginal cost in a few public utilities in an economy in which the other 'optimum' conditions are certainly not fulfilled, and where price is not equal to marginal cost elsewhere. It is indeed exceedingly difficult to imagine a *real* economy in which one could say with any reasonable degree of confidence that if price was made equal to marginal cost then the gainers could compensate the losers. Although, then, there seems to be something in Mr. Hicks's justification, we must conclude that there is certainly not enough in it to justify the extraordinary emphasis which, in recent economic literature, has been laid on the fulfilment of the 'optimum' conditions of production and distribution.

---

[1] My italics.

Whatever the merits of the justification discussed above, in relation to certain kinds of change, there are other kinds of change to which it could not be reasonably applied. For instance, the institution or removal of a general tariff would be a change which could not reasonably be said to form part of a series of changes, the distributional effects of which would cancel out. Taking, then, a change which has to be decided alone on its own merits or demerits, we find ourselves back at the difficulty that such a change cannot properly be said by the new welfare economists to increase economic welfare unless full compensation is paid to everyone, which is impossible and might be undesirable.

Unless we admit the discussion of changes which alter the distribution of real income, and set a valuation on the alteration, we remain within the realm of what Professor Samuelson has called ''twere better' statements. 'Twere better to make the change and compensate all losers, rather than do nothing. The Hicksian justification we discussed above is, in fact, an appeal to the conservative forces of nature to make our compensations for us.

Professor Samuelson seeks to emerge from this realm of useless statements by introducing some mythical personage, who arranges all conceivable configurations of the economic system in order of value. He lays down an economic welfare function. A position of increased welfare is a position which is higher up on the perfectly ordered scale of values belonging to this personage. What this really boils down to is that welfare conclusions cannot be reached without value premisses. The introduction of such a 'policy-maker' is, of course, only an expositional device. It must not be taken to mean that the welfare economist has to wait for someone to lay down value judgements so that he can get on with the job of making deductions. There is no reason why we should not reach welfare conclusions so long as we make clear what our premisses are. Our chief point is that among our premisses there must be some value judgement about the distribution of real income. This remains the case even when we think that distribution can be left to take care of itself. Mr. Hicks, in effect, argued that distributional effects can be neglected. This outburst of classical optimism seems to involve (*a*) the factual statement that such effects will be small, and (*b*) the value judgement that, being small, they ought to be neglected.

---

[1] It should be noted that the advocacy of actual compensation does not (as has been stated) necessarily 'betray' a conservative bias. It may rather 'betray' a bias towards equity. It is quite consistent to hold that the rich ought to be made poorer, and that those rich people who lose as a result of a certain change ought to be compensated. Indeed, this attitude appears to prevail in England at present.

Why are some economists so keen on avoiding obvious value judgements? Presumably because they want to be scientific and objective. If this is their wish, then they must not become welfare economists. They may, of course, confine themselves to the formal logical system which has recently called itself welfare economics, so long as they take care not to give their system a suggestive interpretation. There is little objection to such conclusions as 'x will extend the possibility-locus to the north-east', if it is made clear that this form of words must not be interpreted as 'x will increase welfare'. However, if economists do so confine themselves they are not, of course, being scientific or objective; as pure mathematicians or logicians they live only in the realm of symbols and words.

We have already introduced the value judgement that it is a good thing that an individual be in a chosen position. We now propose that every welfare conclusion must be premissed by a judgement to the effect that the change in question will at least not lead to an unfavourable distribution of income. As an alternative we could use the premiss that a man is happier if in a chosen position, in which case our conclusions about the happiness of the community would require the premiss that the change will not cause a redistribution which would tend to decrease the happiness of the community. We propose to use the former set of premisses, because we are discussing welfare economics, and not happiness economics. The two are not the same thing. The word 'welfare' is at once less descriptive and more recommendatory than the word 'happiness'. It would perhaps have been better if happiness economics had not given place to welfare economics. People would have been more inclined to suspect such a conclusion as 'the happiness of the community would be increased if price were made everywhere equal to marginal cost' than they do the conclusion 'welfare would be increased if . . .'. The difference is terminological, but not *merely* terminological. Terminological differences can be of enormous importance. Happiness conclusions require happiness premisses, and welfare conclusions value premisses.

Having decided to introduce the distribution of real income as a variable, we now have three different kinds of change.[1] First there will be changes which would lead to an unfavourable redistribution.[2] These we can dismiss fairly briefly. They do not satisfy our most important premiss, so it does not matter to us whether the Kaldor–Hicks criterion be satisfied or not. This is not to say that there is anything fundamentally indeterminate about these cases. It might, for instance, happen that a rich person was made so very much better off that someone, who considered that distribution ought to be more equal, would nevertheless judge the change to be good, although a poor man was harmed. Our system cannot, however, say anything about such cases. If the change in distribution is unfavourable, nothing follows. It may, however, be noted that it might be possible in such cases to pay compensation in such a way that there would be no unfavourable distribution effects. In that event compensation ought to be paid, and the case then comes under one of the following two classes.

The second class comprises those changes the distributional effects of which would be thought to be small, such that it is therefore thought that these effects would not be unfavourable. Of such changes as these we can say that they increase welfare if they also satisfy the Kaldor–Hicks criterion. The ground for saying this is not that everyone would be better off (even if every man was an economic man). It is rather that a similar but 'neutral' redistribution of income (e.g. by lump-sum payments) could not make everyone as well off as they would be if the change were carried out. The redistribution by itself would not decrease welfare. Therefore the same redistribution carried out in such a way that the gainers could overcompensate the losers would increase welfare.

Our third class consists of changes which would have a considerable but favourable effect on real income distribution. This class is more complicated than the other two. We can, indeed, say that such changes would increase welfare if they also passed the Kaldor–Hicks criterion. But that is not to say that the change ought to be made, because it might conceivably be possible to increase welfare more by carrying out the real income distribution which the change would cause, in a 'neutral' manner and without making the change. The explanation of this possibility requires a discussion of Professor de Scitovszky's objection to the Kaldor–Hicks criterion.

---

[1] A more formal presentation of some of the argument of the following two pages has been given in my article 'Welfare and Tariffs', *R.E.S.* xvi (2).

[2] It should be noticed that we can lump together questions of justice and welfare, and simply ask whether a certain change in distribution would be a good thing. A change in distribution could be said to increase welfare but be undesirable on grounds of equity. We require as a premiss the judgement that the redistribution would not be bad, taking everything into consideration.

De Scitovszky pointed out that the Kaldor-Hicks criterion for an increase in welfare might give rise to a contradiction.[1] Suppose that a change passed by this criterion is made, no compensation being paid. We may then have a distribution of real income which might differ considerably from the initial distribution. It would then be possible that the reverse change (back to the initial situation) would also be sanctioned by the criterion. In order to get over this difficulty de Scitovszky suggested a double criterion for an increase in welfare. Welfare is increased if (*a*) the gainers could over-compensate the losers, and if (*b*) 'the people who are against the change would be incapable of bribing those in favour to vote against it, without thereby losing more than they would if the change were carried'.[2] The second part of this criterion is the Kaldor–Hicks criterion in reverse. It asks whether, given that every individual were actually in the position he would be in after the change, the gainers by the reverse change could compensate those who would lose by it. If they could do so, then the Scitovszky criterion (as we will call the criterion (*b*) above) is not satisfied.

This double criterion is, formally speaking, an improvement. But we cannot accept it as a criterion of an increase in welfare. If compensation is not paid nothing alters the fact that there would, as a result of the change, be some actual gainers and some actual losers. If compensation is paid the 'contradiction' cannot arise. Nevertheless the Scitovszky amendment is important. We can use it to tell us whether a favourable redistribution of real income, which a certain change would cause, would be better brought about by making the change, or by redistributing money in a 'neutral' manner.

Let us return to our case, where a change passed the Kaldor-Hicks criterion and caused a favourable redistribution. Given the situation after the change it would be possible, if the Scitovszky criterion is not satisfied, further to increase welfare by making the reverse change and compensating the losers. It clearly follows that it would have been better not to have made the change, and instead to have simply redistributed money. If, on the other hand, the Scitovszky criterion were satisfied, it would not matter whether the Kaldor–Hicks criterion were satisfied or not. Redistributing money alone would increase welfare, and if we actually did this, and if the Scitovszky criterion then proved to have been satisfied, it would follow that we could further increase welfare by making the change and compensating

---

[1] T. de Scitovszky, 'A Note on Welfare Propositions in Economics', *R.E.S.*, vol. ix, No. 1.
T. de Scitovszky, 'A Reconsideration of the Theory of Tariffs', ibid., No. 2.

(because given the new distribution of income the gainers could over-compensate the losers if the change were made). It would, in fact, have been better to have made the change. It follows that changes of this third class ought to be made if income redistribution is favourable, and if the Scitovszky criterion is satisfied. In the second class of cases, the two criteria coincide; the Kaldor–Hicks criterion is therefore formally redundant in all cases.

Let us sum up the above arguments. First, two general value premisses are required: (*a*) that an 'individual' is better off if in a chosen position, (*b*) that it is a good thing that one 'individual' should be better off if none are worse off. Secondly, for every particular change we require a judgement as to whether the resultant redistribution of real income would be a good thing or not. If it would be a good thing, or at least not bad, we can say whether the change ought to be made or not. It ought to be made if it would be impossible to make everyone as well off as they would be after the change, by a policy of redistributing money by lump-sum payments.[1]

I must add that the above arguments are formal. I do not want to maintain that the scheme proposed for the foundations of welfare economics is of any use. That depends on how realistic the factual assumptions are which we have to make (e.g. consistency of choice on the part of eternal individuals, absence of dynamic elements, applicability of a marginal analysis, &c.). These we have not the space to discuss here. The question whether our static welfare model is sufficiently realistic to be useful is a very difficult one to decide. It could in principle, but not in practice, be tested. In principle we could find some consistent person who would make the required value judgements, and then we could make the changes which we deduced to be good. If these were the only changes, and if our guinea-pig judged that welfare had indeed increased, then the theory would have worked. In the absence of such verification we can only make untestable guesses as to whether or not the obvious unreality of the axioms is so great that the conclusions are worthless.

Finally, I will briefly give reasons why the present approach is preferred to the 'economic welfare function'[2] approach, from which practical welfare conclusions can only be deduced if an 'ideal' distribution of real income is given. If an 'ideal' real income distribution has to be presupposed, then there are three possibilities:

(*a*) In a democratic society economists assume that whatever the distribution of real income is, it is ideal. This is absurd. It is based on the undemocratic Hegelian view that the government is always right, and on the false idea that distribution is always what the government wants it to be, because if it were not as it wanted it to be, then it would have changed it. This is like saying that I must always want to be in the town in which I am, even if I happen to be motoring rapidly through it.

---

[1] When I say 'ought to be made', I mean, of course, 'ought to be made from the point of view of economic welfare'. By definition, economic welfare can only vary with the amount of economic goods which individuals have.

[2] See Bergson, op. cit.

(*b*) In an absolutist state the economist must assume that whatever is, is right. Welfare economics is the servant of the State, and is denied to those who happen to disagree with the prevailing distribution of wealth. Even given this, it still need not be true that the government always wants exactly what it has in the way of real income distribution.

(*c*) There is the possibility that the ideal distribution can be determined in an objective scientific manner. This fallacious notion is dealt with in Part II below. Thus, rejecting this possibility, it appears that a welfare economics based on an 'ideal' distribution of real income is applicable at best only in a totalitarian state, and not very well even there. We therefore leave open the question whether a change would cause a favourable, unfavourable, or a neutral change in distribution. Welfare economics can then be of some use to those who do not happen to believe that the prevailing distribution is ideal.

We may now draw some conclusions:

1. Welfare economics cannot be objective or scientific. The phrases 'the welfare of society' and 'the happiness of society' are very vague phrases which can find no place in the statement of any general empirical laws.

2. The new welfare economics is not, as usually formulated, a deductive system which draws out conclusions from generally accepted premisses. Either it reaches no conclusions, or it deduces them from unacceptable premisses.

3. As usually formulated the new welfare economics is liable by its use of suggestive emotive language to beg the question of its applicability, and to induce people to accept conclusions about the real world which they might not accept if they understood clearly how they were arrived at.

4. Welfare economics is a logical calculus of doubtful applicability formulated in quasi-ethical terms. The fact that it is a value system is, however, by no means a fundamental objection to it. It might still be useful to anyone who accepts its premisses, and who is prepared to give an answer to the open question concerning real income distribution. If it is useless, that is because its factual assumptions are insufficiently realistic. If they are very unrealistic, then we could only be confident that our criteria would be satisfied in the real world if the theoretical gains (in terms of an ideal system of economic men, firms, &c.) were very large, that is, if the divergences from so-called optimum conditions were very great. But if these divergences must be very great, it is likely that welfare economics can tell us no more than common sense. I am inclined to believe that welfare economics is of little use. Whether I am right or not is anyone's guess.

PART II

The Distribution of Real Income, and Interpersonal Comparisons

In the first part of this article the concept 'the distribution of real income' has been freely used. But obviously its use implies the validity of interpersonal comparisons. Those who 'deny' interpersonal comparisons thus have to claim that 'the distribution of real income' is a value concept, and that one can only speak of the distribution of money income in an empirical manner. This position is taken up by Professor Lerner in the following passage:[1]

> 'This does not mean that the prices will not change and that such changes will not make some people better off and others worse off, even though their money incomes remain the same. It does mean, however, *that we must disregard such changes*[2] when they are the result of changes in demand for the various goods or changes in the supply while the optimum allocation of goods is maintained. If we could say that in the old situation the actual relative well-being of consumers was just what we wanted it to be, then, when prices changed, we would want to make some adjustment in the money incomes, to offset the change in prices, and leave the consumers in the same relative position as before. But we have no way of directly comparing the well-being of different consumers. Our only objective general indication is their money income, and this has not changed. We may have good reason for believing that one consumer is better off than he was before and that another is not so well off as he was before, but we have no more reason for supposing that the old situation is better than the new one (that is, more like the situation we wish to bring about) than for supposing that the new one is better than the old one.'

This passage clearly shows that indifference-curve analysis cannot tell us in what direction the distribution of real income has changed. As a result, the passage is highly paradoxical. Either it is assumed that we cannot ever know when the distribution of real income (or actual well-being) is what we want it to be, or it is assumed that we can know when it is what we want it to be, *although* we cannot tell whether it is becoming more or less like what we want it to be when it is seen to be changing. The latter alternative is absurd, and so it must be supposed that Professor Lerner is denying that one person can ever be said to be richer, economically, than another in any (objective) sense other than that he has a greater money income. This is, as we have said above, the only possible position to take up, if one insists on denying the validity of interpersonal comparisons of utility. But it is a most unconvincing position. It amounts to telling people that they ought to ignore everything but money income (though perhaps capital would be admitted, in so far as it can be reduced to money terms) when they decide what they think the distribution of income ought to be.

---

[1] A. P. Lerner, *The Economics of Control*, p. 24.

[2] My italics.

Professor Lerner suggests that our only objective criterion of the distribution of relative well-being is money income. The word 'objective' may mean many things, but, since Professor Lerner holds that one *cannot ever say* whether one person has more or less satisfaction than another, it appears that he is using 'objective' to mean 'empirical'. Therefore, unless it is held that one has some metaphysical test of relative well-being, it follows that relative well-being cannot mean anything other than relative money income, which cannot, therefore, be said to be a test of relative well-being. What Professor Lerner evidently means, however, is that money income is all we need be interested in, because it can be proved that an equalitarian distribution of money income is more likely than any other to maximize the aggregate amount of satisfaction in the community. It is worth examining how this conclusion is arrived at.

The general position, Professor Lerner maintains, is that, although we can compare individuals to the extent that we can say that their satisfactions are of the same kind, we cannot say how much satisfaction any one individual has, nor can we say whether a small addition to his income will increase his happiness or satisfaction more or less than a similar addition would increase that of another individual. By assuming that satisfactions are of the same kind, Professor Lerner means, evidently, that they may be added. Now most people think that it makes perfectly good sense to say that *A* is more satisfied than *B*, and very little sense to say 'Let us add *A*'s and *B*'s satisfactions'. But there is no need to press this point, because the method by which Professor Lerner gets from this position to his conclusion is invalid. He assumes that because one never has, he believes, any reason whatever to suppose that *A* has more or less capacity for satisfaction than *B*, it is equally probable that a shift of money from *A* to *B*, when they have equal money incomes, would increase or decrease total satisfaction. From this assumption it would follow that, out of a large number of such shifts, 50 per cent. would increase and 50 per cent. would decrease total satisfaction. But from complete ignorance nothing whatever can follow, except complete ignorance. The assumption of equiprobability as a working hypothesis would be legitimate only if there was some conceivable way of testing the conclusion. But, on Professor Lerner's principles, there is no conceivable way of discovering whether the assumption of equiprobability is justified. This step in the argument is, therefore, illegitimate. Professor Lerner's next step is as follows: Out of a large number of such shifts, 50 per cent. would probably increase satisfaction and 50 per cent. would probably decrease it, but, from the assumption of diminishing marginal utility of money to each individual, each loss would be larger than each gain, and, therefore, total satisfaction would probably decrease.

This conclusion does not seem to follow, because each positive shift might represent a large increase in satisfaction, while each negative shift might represent only a small decrease in satisfaction Thus the conclusion does not follow, because each positive shift might be from a person with a very low capacity to one with a very high capacity for satisfaction, while each negative shift might be from a person with a capacity for satisfaction very little lower than the person in whose favour the shift was made. Evidently the argument from total methodological ignorance to equiprobability must be invoked again. We can conclude that any argument which 'denies' interpersonal comparisons, and at the same time asserts that an equalitarian division of money income would probably increase satisfaction, must be wrong.

Thus the contention that we need not worry about anything except money income, because an equalitarian distribution is most likely to produce the maximum amount of happiness, breaks down. Is the contention at all plausible without this support? Assuming an equal distribution of money income, changes in prices would only alter relative well-being in so far as tastes differed. There is some plausibility in assuming that tastes would not differ very much if money incomes were equal, and, therefore, in saying that money incomes (or rather total resources) would be an adequate guide to distribution. But it would still not be very plausible unless tastes were also assumed to be adaptable and price changes small. Moreover, it would still only be true that relative money incomes would be taken as a *guide* to relative real incomes. One would always be ready to abandon the money income criterion if it was observed that some price change made a great difference to some individual. Moreover, as soon as we come to consider an inequalitarian society, the view that only money incomes are important loses all plausibility. No one believes that one must ignore a large change in the price ratio of necessities and luxuries. One would certainly say that the poor had become really better off, or really worse off, in relation to the rich.

The interest that attaches to the study of the distribution of money income is derived from a belief in a correlation between money incomes and real incomes. No interest whatever would attach to comparative money incomes if it was not believed that this correlation existed. It has been shown that no practical welfare judgement can be made without considering the distribution of income. It has also been shown that the distribution of income in which we are interested is the distribution of real income. Since the heterogeneous collection of goods which different people buy may offer no basis for comparison, it is clear that consideration of the distribution of real income involves comparing different people's states of mind or changes in them.

The 'denial' of the possibility of such comparisons is only made by economists. It is not always apparent in what sense they are 'denied'. It is said that they are not objective or scientific. But these words mean very different things. It has also been said that they are value judgements or ethical judgements. But sometimes economists seem to say that a judgement is a value judgement if it is not 'objective'; and, again, it is not clear what the word 'objective' means. In any case, the 'denial', in some sense, of such judgements has resulted in extraordinary confusions in economic theory. It therefore seems worth while to examine the question systematically.

The first position for those who 'deny' interpersonal comparisons must be that other minds are not analysable in behaviouristic terms, and that behaviour is not evidence in favour of the existence of other minds. It must be noted that, if behaviour is accepted as evidence, then we can certainly compare other minds on the basis of such evidence, for, if we have any evidence for other minds, we have overwhelming evidence. Such a procedure would not be in the least non-empirical, and would be, in that sense, objective and scientific. Therefore, on this view, to deny the comparability of other minds is to deny their existence. The only possible alternative suggestion is that we can, by some kind of metaphysical intuition, get to know that other minds exist, but cannot get to know anything else about them, other than the bare fact that they do exist (for if we could get to know anything else about them, we could compare them). How one can know that something exists, without knowing anything else about it, is impossible to imagine. Anyone who holds this view must, nevertheless, admit that people do make interpersonal comparisons. He therefore has to maintain that they are not true or false, because they are not about matters of fact. They are only significant in that they are ethical (value) judgements. This seems to be the position Professor Robbins maintains, as may be seen from the two following quotations: 'I still cannot believe that it is helpful to speak as if interpersonal comparisons of utility rest on scientific foundations—that is, upon observation or introspection',[1] and 'But I still think, when I make interpersonal comparisons, that my judgments are more like judgments of value than judgments of verifiable fact'.[1]

Thus it is held that, when I say '*A* is happier than *B*', I am making a value judgement. We may ask 'What sort of a value judgement?'—perhaps something like '*B* ought to be helped rather than *A*'? In which case, the statement 'You ought to help *B* rather than *A*, because *A* is happier' becomes a tautology. One can only say that it obviously is not a tautology. Also '*A* is happier than *B*' cannot be analysed in the same way as '*A* is a better man than *B*'. By the latter statement a person may intend to convey some fairly definite empirical information. He may himself use such words as 'good' to describe recognizable states of affairs. But there is no generally accepted descriptive usage for such words, that is, they have no general descriptive meaning. On the other hand, words like 'happy' and 'satisfaction' do have more or less definite descriptive meanings. If a person uses the word 'happy' to describe a recognized melancholic, then it would be correct to maintain that he was misusing the word. If, on the other hand, a person says that a man, generally admitted to be bad, is good, then it is incorrect to say that he is literally misusing the word. We might in practice say this, but we should not then be speaking literally, we should rather be saying that we disagreed strongly with his ethical views. We may also notice that Professor Robbins, to be consistent, must maintain that not only interpersonal comparisons of utility, but also of anger, sympathy, kindness, intelligence, and so on, are one and all value judgements, to which no facts have any relevance. We conclude that interpersonal comparisons are not value judgements.

---

[1] L. Robbins, *E.J.*, 1938, pp. 640, 641.

Nevertheless many people quite rightly feel that there is something in what Professor Robbins says. This is because it is true that interpersonal comparisons are not objective and scientific, in other senses of the words. If a number of people have to read a pointer which, it is agreed by all of them, indicates the true weight of an object, then it is not likely that they will arrive at very different results (although they may differ slightly). If a number of people have to decide whether '*A* is happier than *B*', each uses his own tests, giving his own weights to them, and they may reach quite different conclusions. The extent to which they will do so can, however, be exaggerated. There would be some cases in which almost everyone would say that *A* was happier than *B*, but there would be many cases where there would be some disagreement, and where many observers would simply say that they could not tell. In this sense, then, interpersonal comparisons are less scientific and less objective than measuring weights, but it may be noted that it is only a matter of degree.

The fact that interpersonal comparisons are not very objective in this sense is not, however, a ground for 'denying' them. If a person says 'the poor ought to be helped rather than the rich', his statement is quite insignificant unless he can at least sometimes judge 'this man is poorer than that man'. Therefore, a denial of interpersonal comparisons makes complete nonsense of all factual and all general value judgements about income distribution, and consequently makes nonsense of the whole of welfare economics, and much else besides. It must be noted that the statement 'the poor ought to be helped rather than the rich' does not lose significance, because everyone would not agree as to who was poor and who rich.

The bogy of 'interpersonal comparisons' has surely tortured economic theory for long enough. Professor Robbins was right in thinking that welfare economics involved ethics, but wrong in thinking that ethics crept in via interpersonal comparisons. The ethics is there because the welfare terminology is an ethical terminology.

It has been shown that interpersonal comparisons of mental states, or mental characteristics, are essential to judgements of relative economic well-being (for short, in what follows, the word 'wealth' will be used for 'economic well-being') and, therefore, to judgements about the welfare of society. It is apparent, however, that judgements of relative wealth are not direct comparisons of mental states. What, then, are they comparisons of? It might be suggested that they are comparisons of differences in mental states resulting from small changes in money income. Of course, people do not arrive at such judgements by actually going about and offering people small presents to see what difference it makes to them. Nevertheless, whatever evidence is taken into account might be used for estimating the marginal utility of money for different people. A lower marginal utility of money might then be taken as a definition of greater wealth.

The above suggestion is plausible for most cases. People often do make such remarks as 'five shillings a week makes more difference to me than it does to you'. This shows that the idea of comparing marginal utilities of money is a familiar one, which may play a considerable part in judgements of relative wealth. In most cases, it would be paradoxical to say '*A* is wealthier than *B*, but money means more to him'; but not in all cases. A melancholic with hardly any money might, nevertheless, take less delight in a little extra than a man whom most people would want to say was really wealthier.

We may conclude that comparisons of wealth are neither comparisons of money wealth alone, nor are they simply comparisons of states of mind, or of differences in states of mind (although the latter probably play a part). This does not mean that they are non-empirical or meaningless. If I am asked what I mean when I say '*A* is really wealthier than *B*', I can give an account of the evidence which I have considered. This is quite sufficient to make the judgement significant. This evidence would normally include such things as capital, income, prices, and conditions of work, and so on; but it would also include the tastes of the individuals concerned. It is the difference of tastes which makes a comparison of the more objective factors inadequate.

It may be added that people do not claim to make more than very rough comparisons of real wealth or real income. Someone might judge that two people, one a smoker and the other a non-smoker, were economically equal.

Then the price of tobacco might double. Even if he knew this, he might still judge that they were economically equal, although he would admit that one was worse off than before, and the other no better off and no worse off. There is no contradiction in this, because 'economically equal' does not mean anything exact, any more than 'politically equal' or 'socially equal' means anything exact. If the so-called contradiction was pointed out, the person would say 'Oh, by equal I mean more or less equal'. 'Equality' in such a context never means more than 'rough equality'. But if 'economic equality' meant 'having the same money income', a contradiction would arise. For this reason the distribution of income, in so far as it has welfare significance, cannot mean anything very precise.

The fact that the 'distribution of real income' is an imprecise concept is a reason for preferring the type of analysis adopted in the first part of this article. We can inform our 'policy-maker' that the shift in real income distribution would be towards certain social classes and away from others; on the basis of such information he may be prepared to say whether he considers such a change good or bad. We cannot explain exactly what the distribution of real income would be, because we have no language for making such statements. For these reasons a partial analysis which seeks to ask whether one situation is better than another is preferred to the general approach which seeks the optimum by way of an 'economic welfare function', in which all configurations of the economic system are given an ethical valuation. No 'policy-maker' would ever be willing or able to lay down an 'economic welfare function', but he may be willing and able to state in what direction he thinks purchasing power should be shifted. The 'economic welfare function' is a brilliant device which completes the formal mathematical system of 'welfare' economics, but it does not appear to lend itself to interpretation, even in a state with an absolutist government. For instance, the only way a dictator could lay down the ideal real income distribution would be by ostensive definition. He might say, after certain adjustments had been made, 'Now the present distribution of money represents the ideal real income distribution.' His economic advisers would then have the task of keeping those 'optimum' conditions of production and exchange, which could be deduced from his value judgements, in operation without altering the ideal real income distribution. But prices would change, and then how would they know whether the ideal was preserved or not? Evidently it would still be necessary to ask the dictator whether the real income distribution effects of each change would be neutral or not. And if they were not neutral all the difficulties involved in paying compensation would crop up again.

The general conclusion which emerges most clearly from the above discussion is that little can be gained by attempts to be objective and precise about a subject in which people have an ethical interest, and about which they speak in terms which, in so far as they have any descriptive meaning, are essentially vague. In particular, it may be added that our discussion is not meant to imply that the new welfare economics, or any other welfare economics, is *complete* rubbish. It only becomes absurd if it is pushed too hard in the attempt to precisify what cannot be precise, and if perfect scientific objectivity is claimed. Interpreted in a sufficiently vague manner there may well be something in it. It is silly to suppose that a discipline is valueless because it is not scientific. How valuable welfare economics can be as a guide to policy is a question which each individual must decide for himself; but too little has been written which might help one to decide, and too much which serves to impress a decision on one.

ALL SOULS COLLEGE,
OXFORD.

THE ORIGINS OF INDUSTRIAL PEACE: THE CASE OF THE BRITISH BOOT AND SHOE INDUSTRY[1]

By ELIZABETH BRUNNER

THIS is the story of the 1895 dispute in the boot and shoe industry. Relations between the Union and employers in the boot and shoe industry to-day are based on mutual respect and co-operation, and there is a long record of industrial peace, but it was not always so. One of the first things which anyone in the industry will say to the stranger, is that there has been no strike or lock-out since 1895. The conciliation machinery then established has been taken as a model by other trades, and the character of the industry as we know it has developed gradually since then. Yet, although this date is on everybody's lips in the industry, the account of what happened in 1895 had to be pieced together, from contemporary accounts in newspapers and journals, from verbal reports of men who lived through it, from Union history, and from occasional references in secondary authorities.

Fundamentally, the crisis arose in 1895 out of the change in organization that was necessary in passing from the ancient craft to the modern machine industry. The craft of shoemaking, of course, goes back to the earliest times when every farmer knew enough to make rough shoes for himself and his family. In the Middle Ages trades became organized into craft guilds, and the Guild of Cordwainers was one of the earliest. For many years it was nearly all bespoke work, but, towards the end of the handicraft age, shoemakers were also making 'sale' work for stock. The giving-out of work then came naturally. The capitalist entrepreneur came between the shoemaker and his public, the entrepreneur undertaking the risks and the profits of marketing and giving out the work to be made by shoemakers working in their own homes and usually helped by members of their families and perhaps boy labour. With the widening market this was the natural organization, for with an increased scale of operations the profit margin could be narrower on each pair of 'sale' work than on 'custom' work, and labour costs were lowered by the subdivision of operations. Some of the older people connected with the industry can still recall the later phases of this organization. Many businesses still surviving date their inception from that period, for it involved the establishing of a 'central shop', which later developed into the factory. The central shop originated in the entrepreneur's desire to have more control over the people working for him, because he found difficulty in getting work out exactly to specification or on time. Shoemakers were then known proverbially as a wild and undisciplined race who, as soon as they had got their money for the week, spent it all again in the excesses of 'Saint Monday'. This is a contemporary description: 'Shoemakers work late on Saturday night, often lie in bed all Sunday morning, lounge in listlessness during the afternoon, drink all Monday, are sick and taking physic on Tuesday, and return to work on Wednesday.'[1]

---

[1] This article is a by-product of work now being done at Nuffield College. In the course of an inquiry, financed through the generosity of the late Mr. Samuel Courtauld, into 'The relative efficiency of large- and small-scale business', a special study has been made of the boot and shoe industry.

In the central shop, generally, the uppers were cut and they were then sent out to be closed in the workers' homes; when closed, they were returned to the shop and given out again with the equivalent number of soles to the makers, who would last them and return them, and they would be again sent out to be finished. They were then inspected, packed, and dispatched from the centre. A new race of workmen grew up under this system who only performed a few operations, and there was no longer any need for the apprentice who learnt to make a shoe all through. It can be readily seen that, with the advent of machinery, this system was easily adapted to exploit the new inventions, more and more work was drawn indoors, and the central shop grew into the modern factory.

The system itself was easily adapted, but were the men? The men were certainly vocal, for they were relatively well organized. The Amalgamated Society of Cordwainers dates from the eighteenth century; it was reformed in 1862, and, in 1874, took the name of the Amalgamated Society of Boot and Shoe Makers, when the men in the machine trade split off to form their own National Union of Operative Boot and Shoe Rivetters and Finishers, later the National Union of Boot and Shoe Operatives.[2]

At the end of the 1850's the sewing-machine began to be applied to boot and shoe work, and between 1857 and 1874 the invention and introduction of boot and shoe machinery led to a complete trade revolution which was to be consolidated in the next few decades. The men at the point of contact at first bitterly resented the change and there was a lot of unreasoned opposition to the introduction of machinery. It was not only that the men disliked the revolution in methods and skill, but they feared the introduction of a factory system which was abhorrent to them. They had been used to a great deal of independence, and indeed they did not settle down easily in the first years of factory life, although such life was evidently more casual than it would be in a modern factory. A bookseller in Northampton, now in his seventies, told us that, as a lad, he had been in a boot factory, but that because of the unrest of this time he had left the trade. He described how someone might come round the factory and say there was a cricket match at Irthlingborough, and those who wanted to play cricket would simply down tools and go off for the afternoon. Thus the change was eased for the men, but it would, of course, be a source of irritation to the employers.

---

[1] A cutting from an Edinburgh magazine, entitled 'Working Classes of Edinburgh', quoted by Constance Calver in a chapter on the Boot and Shoe Trade in *Seasonal Trades*. Incidentally, it is amusing to note that these habits were known to the employers as 'Black Monday', and in Czechoslovakia, shoemakers being the same the world over, they kept 'Blue Monday'. Politically, also, shoemakers seem to have the same sort of reputation. Constance Calver tells the following story: 'It is related that a deputation of working men waited on Sir Robert Peel to lay the wants of the trade societies before him. The two selected to speak were shoemakers. "How is it," said Sir Robert, "that you shoemakers are found in every movement? If there is a plot, or conspiracy, or insurrection, or political movement, I always find there is a shoemaker in the fray."' Even in 1945 we were reminded that Northampton was the town that elected Bradlaugh, and that shoemakers were 'a terribly radical lot'! About the time of the election of Bradlaugh, indeed, the Church was moved to set up a mission hall in Northampton to try and convert these atheists; it later became offices of the trade union.

[2] The N.U.B.S.O. covers the whole country with the exception of the isolated Rossendale Valley, which has its own union (and the manufacturers have their own federation).

The policy of the Cordwainers' Union, however, was very far-sighted. They avoided the mistake of the handloom weavers, for instance, who tried to compete with the new textile machinery and whose craft vanished in the misery of lowered wages and lowered quality. They 'frankly abandoned to the machine any part of the trade within its scope, and concentrated on maintaining and differentiating the peculiar qualities of their own special article'.[1] The Union resolutely refused to oppose the introduction of machinery, but, at the same time, refused to accept any cuts in the established prices for handwork. Moreover, it set about organizing the new factory workers so that it largely avoided the unfair competition of unorganized women and boys employed at a rate below a living wage. The factory workers' organization thus grew up inside that of the handicraftsmen until, in 1874, it was found necessary for it to break off. The handicraftsmen did not oppose, they even encouraged, the separation, and the two unions since then have remained on friendly terms, refusing to allow themselves to compete or to be played off against one another. There *was* a sweated trade—indeed there could have been a 'Song of the Slipper' to go with the 'Song of the Shirt' in the East End of London—but it was not the skilled handsewn craftsmen who were sweated or the new factory workers. The sweated trade was most prevalent before the wider spread of the factory system; it applied to the same class of work, and employed intense division of labour. It could, therefore, employ some low-quality workmanship, and, to a large extent, pegs and rivets were substituted for stitching in the production of inferior articles. Because of the lack of control over outdoor wages and conditions, the Union fought to get all outdoor lasting and finishing brought within the factory. This was achieved largely by 1891, but it must not be forgotten that in some sections of the industry traces of the old methods remained, even till the 1930's.

---

[1] *Industrial Democracy*, by Sidney and Beatrice Webb.

Mr. (now Sir George) Chester, General Secretary of the National Union of Boot and Shoe Operatives, explains the breakaway of the machine operatives' Union thus: 'The Association (i.e. the Amalgamated Society of Boot and Shoe Makers) had proved incapable of combating the employers' attacks on the local Unions or coping with the changing industrial conditions';[1] and so, by implication, the new Union was to be much more militant, as indeed it was. In its early years there were constant strikes and disputes waged with somewhat inadequate numerical and financial resources, until the Union won through to a position of considerable power, round about 1890. The following membership figures may be useful, as an epitome of the Union's history:

1874	4,204	1896	36,385
1878	5,084	1897	35,598
1888	11,410	1898	32,084
1889	13,760	1899	27,993
1890	23,459	1900	27,960
1891	30,046	1910	30,197
1892	42,524	1918	83,017
1893	41,274	1924	76,100
1894	43,955	1939	84,198
1895	40,720		

The years 1888–92 were a period of rapid growth for all trade unions, and a period of good trade. The slight falling off in membership figures in 1893 reflects a falling off in trade. 1895 was the year of the great lock-out, and the downward trend is clear from that date. Membership figures were perhaps sustained in 1896–7 by the recovery of trade, but 1898 was the year of a long-drawn-out strike in Norwich, where the Union was decisively defeated. It was not until 1900 that the slide was stopped, and the Union set about consolidating its position, reconstructing its constitution, and appointing full-time branch officers, &c. It must be remembered, however, that a slight fall in actual numbers is consistent with a small proportionate increase in Union membership, for the numbers of those looking to the industry for a livelihood appear to have fallen gradually but fairly continuously since 1891 (or earlier) to the present day. The 1914–18 war, of course, meant a tremendous leap forward in trade-union organization, here as in other industries, but the peak membership figures were not sustained, and there was a constant sporadic fight to get the smaller centres fully organized, on both the employers' and the workers' sides, an aim which was, on the whole, achieved by the outbreak of war in 1939.

The events of 1895 have affected the course of the industry profoundly. To an outsider, the immediate causes of the dispute must have seemed quite trivial in proportion to their effects, and the confusion as to motive is apparent in the nomenclature—the dispute being termed either a strike or a lock-out. The truth was expressed by a contemporary writer,[1] when he said it was 'neither a strike nor a lock-out but an industrial war'. The real causes were not to be found in the surface quarrels which were picked, but sprang from the deep malaise of an industry in transition. Each opponent wished for a 'show-down' and a final test of the other's strength.

---

[1] Chapter on the boot and shoe operatives in *British Trade Unionism Today* (ed. Cole).

Other industries had already had their struggles over the Industrial Revolution, but the boot and shoe industry was later than most in changing over from the old handcraft to the modern machine trade, which perhaps made the struggle all the more bitter. What was in dispute was not so much the actual introduction of machinery—that was by then accepted—as the conditions under which that machinery was to be operated. The men, for as long as they could remember, had been used to work on a piece-work basis, and the price-lists for handwork and the early factory work had been achieved by great efforts and were recognized as fair. But, in the new situation in which they found themselves, their hard-won price-lists were useless. They had to go on to time-rates, and the employers resolutely refused to consider fixing new piece-work price-lists. Here there was a fundamental cleavage. The men not only felt cheated of a traditional right, but, under the new conditions, piece-work was the only way of maintaining the workers' position. In an industry, which was being 'incessantly revolutionized by an almost perpetual stream of new inventions and new applications of the old machines', it was easy for the workers to be driven until they were doing more and more work for the same payment. Further, the introduction of the team system, with subdivision of labour, an organizational change, meant that the pace of work could be forced up to keep time with that of a leader. The only way the men could see of improving their position was deliberately to restrict output on the new machines, and men were actually fined from 2*s*. 6*d*. to 20*s*. by the Union for producing more than a certain output.

The men had other grievances too, connected with the new machinery—in particular, 'basket' work, and the employment of a high percentage of boys on the new subdivided operations. To take the latter first, the objection to boy labour was not to any form of apprenticeship in which the boys might start on simple operations and gradually work up, but to the setting aside of operations permanently as boys' work in an attempt to dispense with as much skilled and adult labour as possible. It was blind-alley work for the boys, and, of course, it was another instance of 'unfair' competition leading in the end to a depressing of the standard rate for adults, and so was vigorously resisted by the Union, who asked for the limitation of one boy to five men.

---

[1] W. B. Hoffman, 'The Late Boot War', *Economic Journal*, June 1895.

It was the same sort of reasons that made the Union object to 'basket' work, i.e. the sending out of work from the towns to the country villages, where the 'statement' was inoperative, or sending work from a district with high wages to one with a lower statement. (In London, this took the form of a reversion to homework.) The Union insisted that if a manufacturer had agreed to a certain statement he must pay the wages accordingly, though they had no objection to his setting up a new factory in a country district. The objection of the Union to homework, of course, helped on the revolution in the trade.

The manufacturers, on the other hand, were anxious to exploit the benefits of machinery, because they found themselves outpaced by the American factories. America had managed the change-over from the handicraft to the factory system some years earlier and was now technically far in advance of our own country, and exporting to it. The manufacturers said they could not meet American competition so long as the Union interfered in the internal management of their factories. To the charge of 'exploitation' from the work-people they opposed the charge of 'obstruction'. They claimed that the men's restrictive practices reduced output by as much as a third in many factories. The editor of the employers' journal[1] thus expressed it:

> 'The men working the machines exercise all their ingenuity in making machine-work as expensive as hand-labour. There exists among workmen what amounts to a tacit understanding that only so much work shall be done within a certain time, and, no matter what machines are introduced, the men conspire to prevent any saving being effected by their aid. It is of no use to mince words. The Unions are engaged in a gigantic conspiracy to hinder and retard the development of labour-saving appliances in this country. The action of their members in failing to exercise due diligence in working new machines is equivalent to absolute dishonesty. It is, indeed, positively painful to any-one who has been accustomed to see, for example, finishing machinery running in American factories, to watch English operatives using the same machines. In America the men *work*, they run the machines to their utmost capacity, and vie with each other in their endeavour to get through as much work as possible. But in an English factory they seem to loaf away their time in a manner which is perfectly exasperating. If they run a machine for five minutes at full speed, they seem to think it necessary to stop it and see that no breakage has occurred. Then they walk about the shop, and borrow an oil-can or a spanner, wherewith to do some totally unnecessary thing. This occupies anywhere from five minutes to an hour, and then the machine is run on again for a few minutes. . . .'

Again, E. A. Pratt, writing in 1904 a general diatribe against ca' canny in British industry,[2] comments on the same 'art of measuring up to a nicety the quantity of work to be done' in shoe firms, and if any operative did more his life might be made intolerable. Pratt adds that 'the adoption of this short-sighted and retrogressive policy by the Union, and the consequent check to the development of industry in the trade, were the chief causes of the lock-out in 1895'. The manufacturers objected to the Union's demand for a minimum wage, too, on the grounds that it would mean 'lower productivity'.

---

[1] *The Shoe and Leather Record*, 19 Feb. 1892, quoted by S. and B. Webb, *Industrial Democracy*.

[2] *Trade Unionism and British Industry*, by E. A. Pratt.

Each side claimed that all the advantages of the introduction of machinery should belong to it. This was why the manufacturers shortsightedly clung to day-work. They thought that, if they conceded a piece-work scale, the men would increase their skill and quickness until they were perhaps earning two or three times the time-rate. The manufacturers could not bear with equanimity the prospect of the men 'making too much', even though there was no question of this rise in wages meaning a corresponding rise in the cost of production per pair, which would, in fact, probably be lowered considerably by the resulting increase in output. The men were equally unreasonable, though in the opposite direction. They could not bear the thought of the increased reward going to anyone else, and they drew up their proposed piece-work price-lists on the basis that the cost of each boot to the manufacturer should be no less than it was before. Thus, instead of frankly demanding an increase in wages (which, granted the increased productivity of the industry through the new machinery, they should have got) they resisted any departure from the old statement prices as though it were a positive reduction, and this the employers resented.

It will be seen already, from this description of the attitudes taken up, that the inherent difficulties of an industry in transition were exacerbated by bad relations between the employers and workmen. By the 1890's, as we have seen, the Union was in a very powerful position and the manufacturers distrusted its growing power. They came up against it, more particularly, in what they considered unwarranted interference in the internal organization of factories, a matter which the manufacturers considered was solely the responsibility of the management. More generally, there was also political antagonism. The Union was imbued with the aggressive spirit of the new unionism of the time, and its accompanying drive towards socialism. In 1894 the Union's rules were revised in a socialistic direction, including amongst a long list of objects 'the introduction of industrial co-operation in the trade' and the advocacy of 'the nationalization of the land and implements of production and their distribution'.[1] To bring in politics must have seemed to the manufacturers an additional and unnecessary provocation.

---

[1] Ref. Charles Booth, *Life and Labour of the People*, vol. vii, chapter on the Boot and Shoe Makers, by G. E. Arkell.

While the shoemakers were renowned for their turbulence and lack of discipline, and their Union was avowedly militant, the employers were themselves drawn from the ranks of shoemakers, so had the same belligerent character. They were only the first generation of capitalist entrepreneurs. As might have been expected, many of the employers had less sympathy with the point of view of the workmen than have those with a tradition of leadership over several generations. The Manufacturers' Federation, moreover, had come into being (in 1891) because the local associations felt at a disadvantage in dealing with the Union, and so its main purpose was to fight the latter. That bad relations were the cause of the matter can be seen in that people who lived through the time have now quite forgotten the *causes* of the strike. But everyone remembers the bitter feeling engendered and that it was 'a very bad time'. The only explanation put forward to us by an employer of that time was that it was all due to 'awkwardness and devilment' on both sides. What is remembered is the workers' processions led by home-made bands; and a manufacturer recalls rejected loaves of bread being thrown into his garden by the strikers, when he was a child.

Relations between the two sides were regulated by conciliation machinery which was, for its time, remarkably well developed. It is worth stressing this early establishment of a policy which is such a noteworthy feature of the industry's history, both for its intrinsic interest and because it did mean that there was some tradition of peaceful arbitration to build on in the settlement of the great 1895 dispute. The first General Secretary of the Union, Mr. Thomas Smith, saw the value of conciliation and arbitration, which he advocated as early as 1874, and in February 1877 the first case of its kind was submitted to an independent referee during a four days' sitting of an Arbitration Court.[1] In every important centre of the trade from 1875 onwards there was established a local board of conciliation and arbitration, formed of an equal number of manufacturers and trade unionists, elected locally, to which had to be referred 'every question, or aspect of a question, affecting the relations of employers and workmen individually or collectively'. In the event of their disagreeing, the matter was referred to an independent umpire. Similarly, later, from 1892 onwards, national questions affecting the whole industry were dealt with by a National Conference, consisting again of equal numbers of members of the Federated Association of Boot and Shoe Manufacturers of Great Britain and of the National Union of Boot and Shoe Operatives; and again, if there was no agreement, the matter was referred to an umpire, Sir Henry (later Lord) James, at one time Attorney-General.

---

[1] Ref. Sir George Chester, 'The Boot and Shoe Operatives', chapter in *British Trade Unionism Today*, ed. by G. D. H. Cole.

Theoretically this was doubtless excellent, but, as the Webbs put it, 'the boards have been the cause of endless friction, discontent, and waste of energy among workmen and employers alike'. Unsatisfactory to both sides, relations got more and more strained and 1895 was the breaking-point. The troubled years of 1891–4 saw no fewer than seven dissolutions of the local boards. The men objected to them principally on the grounds of the extreme slowness with which they did their work. When any general agreement had been concluded, it might be months before the local interpretation of it was decided by the various boards. Then, for technical reasons, the umpire was usually a manufacturer from another district, and if the award went against them the men were quick to suspect bias. Given this lack of faith in the boards and a turbulent element among the men not really controlled by the Union, it was not surprising that there were frequent revolts against the working of the local boards. But it is significant that the awards they flouted nearly always concerned the newly introduced machinery.

The manufacturers too were dissatisfied. The boards had complete authority to settle any wage questions, &c., without reference to any central control. If they were not to let slip important questions and find themselves committed to some disputed principle before they were aware, the big manufacturers had to spend a considerable amount of time which they could ill spare attending to the laborious and incessant work of the boards, when much of it was only of minor importance. The real trouble was that the manufacturers, in setting up this conciliation machinery, had committed themselves to something they never intended. In all matters of collective bargaining there are two distinct aspects, that is, the making of a new bargain and the interpretation of the terms of an existing one. The boot and shoe conciliation machinery did not work smoothly as long as this distinction was not perceived or maintained. The manufacturers obviously intended the boards to deal with interpretation questions only, i.e. 'not the amount per week which the workmen in any particular establishment should be permitted to earn—for that has, in principle, already been settled—but the rate at which, under the actual conditions of that establishment, and the class of goods in question, the piece-work price must be computed in order that the average earnings of a particular section of workmen shall amount to no more and no less than the agreed standard'.[1] That is essentially a question of fact and the ascertainment of facts as precisely as possible. The questions of principle, such as the standard wage-rate or the question of boy labour, they had not contemplated being decided by these boards, although they had agreed to a rubric which certainly included them. On these questions of principle, the two sides never could agree, as long as the National Conference simply meant that each stated their own viewpoint and then the umpire gave an award by arbitration. This put Lord James in the position of supreme legislator for the shoe trade, which encroached on the manufacturers' position. Moreover, the manufacturers were never free from pressure for an increase in wages—the question could be brought up every other month at the local boards, and the National Conference met five times in the three years 1892–4; and they complained that the men did not keep agreements anyway. Things came to a head when the British market was increasingly entered by cheap machine-made shoes from America—still recalled as the 'American invasion'—and the manufacturers realized acutely their technical inferiority in the matter of machinery. Yet every improvement was made the occasion of a long wrangle between employer and workmen, still further retarding efficiency.

---

[1] *Industrial Democracy*, by Sidney and Beatrice Webb.

The manufacturers struck in the autumn of 1894, choosing a period of slack trade. They submitted to the Union seven proposals (to be known as the 'Seven Commandments'), the object of which was described as, 'by affirming certain principles to restrict the operation of the Arbitration Boards to the work for which they were originally intended'. The proposals expressly reserved certain matters as outside the machinery of arbitration. These included: any change of wages within the next two years or within two years of the date of any subsequent award; the introduction at present of piece-work for machine-work (and, when it was introduced, they wished it to be based on the average day-work wages earned); any restriction on the employers' sole right to regulate the internal management of their factories, in paying either piece- or day-work wages, in introducing machinery, in having their work done where they pleased (i.e. 'basket' work), or in employing whom they pleased; any restriction of output by the Union. From then on a big strike or lock-out was inevitable, and the Union began collecting funds for the battle, but, at the same time, it did not want to join issue just yet. The Federation had demanded the acceptance of their proposals within eleven days, but it was not until the end of January that the Union rejected them as 'illegal, unjust, unworkable, and, therefore, impracticable', though, at the same time, they suggested arbitration. The employers meanwhile had withdrawn from the National Conference and had instructed their members on the local boards to refuse to discuss matters covered by the seven proposals. They now again refused arbitration. The Union, therefore, at the end of February, at the approach of the busy season of spring trade, challenged the manufacturers' position on the grounds of the powers of the local boards. At Northampton the Union applied for a full meeting of the local Arbitration Board to consider their request for an increase in the minimum wage of clickers and pressmen, but the manufacturers refused, on the grounds that it contravened the first of the Seven Commandments (i.e. that there should be no change in agreed wage-rates for two years). Again, at Leicester, besides raising the question of an advance for clickers and pressmen, the operatives put forward a motion 'that this Board decides that all work cut in Leicester shall be made and finished in Leicester, and paid for in accordance with the prices and conditions at present in operation at Leicester', and the manufacturers again declined to discuss it, as being outside the jurisdiction of the Board. The Union saw they had no course open but to strike, and notices were served on three firms in Northampton and six in Leicester; on 6 March the Federation resolved to meet this strike by a general lock-out.

The actual stoppage, then, came on side-issues of minor importance. The relevance of the 'basket' work question has been explained. The issue of an advance in wages for clickers and pressmen seems to have been drawn in almost at the last minute, as a matter of internal Union politics. These classes of workmen had only lately joined the Union, which had been founded primarily for rivetters and finishers, and the Union wished to show that it could do something for them; also, perhaps, it was a way of ensuring their solidarity with the older members in the coming strike (clickers, at any rate, were still handcraftsmen and unaffected by the new machinery).

The way in which the matter came about shows Leicester, the headquarters of the Union,[1] acting as the most influential district at that time. Altogether, about 46,000 operatives were affected, in Leicestershire, Northamptonshire, London, Leeds, Bristol and Kingswood, &c. But Scotland and the English centres not affiliated with the Federation—Stafford, Norwich, Stone, Wellingborough, and Hinkley—were not affected.

The struggle was long, because the Federation persistently refused to agree to arbitration, firstly because they did not consider the questions at issue as fit subjects for arbitration, and secondly because they had doubts as to the Union's faithfully carrying out any award that might be given.

Finally, through the good offices of Sir Courtenay Boyle, Permanent Secretary of the Board of Trade, the two sides came to an agreement. A series of informal conferences was held at the Board of Trade, culminating in a six-hour conference on 19 April, when the two sides of the industry were in separate rooms, and Sir Courtenay Boyle and other officials had to go from one to the other. The Terms of Settlement were signed then which, with slight revision, have regulated labour relations in the industry ever since.

---

[1] In 1898, when it came to the turn of Norwich to strike, and the Union was ignominiously defeated, it appears that the strike was called, and hopelessly prolonged, by the central executive from Leicester, the local secretary of the Union being all along opposed to the strike. In 1895, however, the Great Strike was a truly national affair, and there was a solid backing to Leicester's leadership. Leicester only came to its position of importance with the rise of the machine industry, when the strikes against the introduction of sewing-machines, &c., in the old-established centres like Northampton and Stafford caused the industry to become more widespread.

The text of the terms of settlement falls into two halves, one dealing with piece-work statements and one with Boards of Arbitration. It was agreed, on the first, that a piece-work statement, or statements, for lasting and finishing was desirable (a point for the Union), but such statements were to be based on 'the actual capacity of an average workman'. (This was directly opposed to the attitude previously taken up by the Union, who were afraid, anyway, that the pacemaker would be taken as the 'average'.) Joint committees were set up in connexion with these piece-work statements, one to establish the national principles, and others to apply them locally. It was agreed on the second that the local Boards of Arbitration should be reconstituted and their rules revised and that meanwhile 'only questions of classification and other minor local questions not involving matters of principle' were to be entertained. Further, certain subjects were laid down specifically as outside the function of the boards, on the lines wanted by the manufacturers. Finally, a trust deed was established to provide a financial guarantee for the carrying out of agreements—in particular, of the overriding proviso that neither side should enter into any strike or lock-out. The Federation and Union each deposited £1,000 with the Trustees, and, in the event of one of them being fined for breach of agreement, the funds had to be made up again by the guilty party, so that the fund was always in being to be called upon.

In many ways the terms of settlement were vague and, so, unsatisfactory. No provision was made for a National Conference, and the fundamental distinction between the two types of questions—those of interpretation and those breaking new ground—was not clearly drawn. Each side went back to its members and announced that it had won the victory, whereas it was obvious that they both could not have done so. In fact, the Union had had to surrender a good deal of ground, and in later years they confessed to having signed it only under duress. There were stormy meetings with the men, who declared that the employers had gained 5½ out of the 7 Commandments. But a good deal turned on how the terms were interpreted, and here the standing of Sir Courtenay Boyle served the trade well. There was some trouble and much ill feeling when almost immediately the manufacturers tried to emphasize their 'just rights' that they had won in the management of their own factories, but wiser counsels prevailed. The National Conference was re-established and took place every two years, and this time it was clearly designed to deal with questions of principle, which could thus be raised only then, and the independent chairman was there simply to help the two sides to reach agreement, to act as conciliator rather than arbitrator. The local boards were confined strictly to matters of interpretation and local affairs. A large part of their success is due to the fact that they can and do act promptly, thus dealing with grievances on the spot quickly, before they can spread or accumulate.

The settlement of 1895 did, in fact, work, in spite of all gloomy prognostications to the contrary. The trouble in 1898 at Norwich (a district which did not come out in 1895) was the last major strike. There have, of course, been disputes in individual factories, when the men have struck, but they have not spread to other factories,[1] and they have seldom been allowed to extend beyond the permitted three days,[2] after which time money would be forfeit. For a long time feeling was bitter in the industry, but the physical restraint on strikes and lock-outs and the continuous contact between employers and men's representatives in the detailed settlement of piece-work prices and conditions of work have led to a gradual change in mental attitudes. In this development the Union has played a special part. A personal impression is that they have done much to educate the manufacturers to appreciate the value of the machinery which was jointly set up. Now both sides of the industry look back to the settlement of 1895 with pride, although at the time extremists on both sides regarded it as a betrayal.

NUFFIELD COLLEGE,
OXFORD.

---

[1] The Chesham strike in 1922 is in a class by itself. The Chesham manufacturers withdrew from the Federation and refused to observe the conditions of the National Agreement. The Union, therefore, withdrew all operatives, except from the Chesham Co-operative Society, and they had the full support of the Federation, who voted to supplement the men's strike pay for those who had children under 14.

[2] Originally ten days; altered to three in 1910.

A NOTE ON CURRENT MARXIST DEFINITIONS OF THE NATIONAL INCOME

By DUDLEY SEERS

ON 18 February 1948 a conference was held between representatives of the Polish Workers' Party and the Polish Socialist Party to discuss the technique of planning. The conference was opened by Mr. Minc, Minister of Industry and Trade, who read a statement on the place of national income estimates in economic planning, and suggested alterations in national income definitions which would take them far from those current in the rest of the world, but put them apparently on a similar basis to those used in Russia. Extracts from this statement have been translated into English and distributed by the Polish Embassy.[1] The statement has also been widely discussed in Eastern Europe, and it seems likely that the national income estimates in that part of the world will conform to the definitions laid down by Mr. Minc. The status of Mr. Minc, his quotations from Marx, and his references to Soviet experience ensure a very serious consideration by economists in the new régimes of Eastern Europe. A similar set of definitions is favoured by Mr. Kochanovskov, a Czech economist.[2] It is therefore worth while examining Mr. Minc's statement to see what sort of national income estimates he proposes should be adopted, and what their use would be. In the course of this note some criticisms of the use of Marx's economic categories for administrative economics will be developed.

Mr. Minc first considers 'the fundamental principles of the Marxist theory of national income'. Following the usual Marxist symbolism, C is the depreciation and material use; V is the variable capital or labour power used; S is the 'surplus value' produced by that labour. Two aggregates are then distinguished: Total Product is $C+V+S$; and National Income is $V+S$. This second aggregate is 'the sum of new values created in a given year as a result of work expended during that year', and 'work' is immediately qualified to mean 'work which is productive in the sense of Marxist political economy', i.e. 'work which is direct harnessing of the goods of nature'. Book-keeping, art, education, medical services, defence, finance, science, public administration, consumers' services, and trade (other than 'functions which are the direct extension of material production in the sphere of distribution') are

---

[1] The most important part of Mr. Minc's memorandum is reprinted in the *Communist Review*, Sept. 1948.

[2] *Sociálni Revue*, published by the Czechoslovak Ministry of Social Welfare, 1948, Nos. 8–9, 9–10, 10–11. I am very grateful to Dr. Stadler, First Secretary to the Czechoslovak Embassy in London, who translated these articles for me.

'non-productive'.[1] Transport, storage, and communication are 'branches of material production'.

This division, which appears somewhat arbitrary, is based on certain distinctions in volumes ii and iii of *Capital*. According to Marx, the function of commerce is 'the transformation of capital from commodities into money and from money into commodities', and 'these transformations constitute merely a conversion of the same value from one form into another'. 'This work . . . does not create value any more than the work done in a civil process increases the value of the object of contention.'[2] As for book-keeping, 'it is the same condition of things which we observed in the case of the time of sale and purchase'.[3] All such 'expenses of circulation, which are due to a mere change of form in circulation, ideally speaking, do not enter into the value of the commodities'.[4]

'Not so', Marx goes on, 'the expenses of circulation which we shall consider now. They may arise from processes of production, which are continued only in circulation. . . . Or, on the other hand, they may represent from the standpoint of society mere unproductive expenses of subjective or materialised labour, while for this very reason they may become productive of value for the individual capitalist, by making an addition to the price of his commodities.' Among these latter 'dead expenses' is storage, by which the diminution of value through deterioration 'is restricted',[5] but commodity storage creates no value if it results merely from a delay in circulation.[6] Transport also creates value since 'the use-value of things has no existence except in consumption, and this may necessitate a change in place on the part of the product, in other words it may require the additional process of production of the transportation industry'.[7]

It is in fact not at all clear what is the criterion of 'productive work.'

---

[1] One point of distinction from Western treatment is brought out sharply by Mr. Kochanovskov when he states that the national income will be no higher if everyone is given a modern hair style by hairdressers. 'Budeme-lí míti moderní účesy, tím zřejmě národní důchod se nezvětší, a tudíž se nesmí příjmy závodu takového druhu započítávati do národního důchodu; patří do sféry konsumu' (*Sociální Revue*, 1948, No. 10–11, p. 287).

[2] *Capital* (Charles Kerr ed., 1909), vol. ii, pp. 147–8 (beginning of chap. vi).

[3] Ibid., p. 151. Mr. Mine states: 'This is only a reflection of the real process and only an ideal generalisation of the real process' (of production)—hence it is 'non-productive'. Mr. Kochanovskov also considers supervisory labour 'non-productive' (*Sociální Revue*, 1948, No. 10–11, p. 286). Some services connected with industry are considered 'productive' by Mr. Kochanovskov, however, because of their relation to industrial output. Research financed by industry, for example, should be included in the national income, he considers, while that financed by the government should not, and he quotes as an example the sweeping of chimneys, which he would include in national income if done for industry, but not if done for consumers.

[4] Ibid., pp. 154–5.

[5] Ibid., p. 158. 'The book-keeping incidental to this process does not influence the use-values . . . .'

[6] 'Inasmuch, therefore, as the formation of a supply involves a delay in the circulation, the expenses caused thereby do not add anything to the value of the commodities' (ibid., p. 165). This is an 'abnormal' stock or 'supply'. Marx admits that 'the normal and abnormal form of the supply cannot be distinguished externally'.

[7] Ibid., p. 170.

As Dr. Paul Sweezy says, with some understatement: 'Marx's treatment of these [merchant's] expenses is not altogether unambiguous; the relevant passages have the earmarks of a rough first draft . . . .'[1] The plea of the necessity of transport and storage applies also to commerce and bookkeeping, which too could be described, as 'the direct extension of material production into the sphere of distribution'. Mr. Minc, it should be noted, specifically rejects such criteria as the manual nature of the work, the 'materialization' of the work in the product, and the product's social use. But his own criterion of 'harnessing the goods of nature' can scarcely be termed precise. It is not at all clear how storage and communication do this, but defence and medical services do not, nor does the other criterion he uses as synonymous, 'reflecting the relations of man to nature', appear any more definite.[2] Dr. Sweezy states flatly that 'this principle [of unproductive commerce] is perfectly plain to the individual industrial capitalist, who knows very well that an increase in the costs of buying and selling, other things remaining equal, does not raise the value of his products but instead reduces his profit',[3] but this of course applies also to transport, and indeed to all his other costs, 'productive' or otherwise. The implication of Dr. Sweezy's statement is that commercial expenses cannot be 'passed on' to the consumer, but while this may be true for individual capitalists, it is not true for an industry as a whole—and the case of commerce seems anyway to be no different from that of transport in this respect. They both increase with the scale of capitalist output and indeed are necessary for this increase. Marxists might well claim that large sections of capitalist commerce have no justification from a broad social viewpoint, and many non-Marxists would agree with them, but this introduces the concept of social usefulness into the definition of what is 'productive' work, a concept which Mr. Minc explicitly rejects and which is incompatible anyway with the division of output made by both Marx and himself, a division which appears to be basically philosophic rather than economic.[4]

This distinction between 'non-productive' and 'productive' work led Marx in fact into some theoretical difficulty when he came to explain how 'non-productive' labour could be exploited. His solution was that the commercial capitalist receives part of the surplus of the industrial capitalist, and 'the unpaid [i.e. surplus] labour of his clerks, while it does not create any surplus value, at least appropriates surplus value for him, which amounts to the same thing so far as results on his capital go'.[1] Perhaps this is formally adequate, but it is certainly clumsy.

---

[1] *Theory of Capitalist Development*, p. 279.

[2] Mr. Kochanovskov believes that we must look to social relations to help us decide what is productive—particularly whether work creates surplus value (*Sociální Revue*, 1948, No. 10–11, p. 285).

[3] Ibid., p. 278.

[4] Mr. Kochanovskov draws special attention to the existence of a category of work which is 'non-productive' but socially necessary.

Marx could quite easily have treated all employed labour, or at least all labour employed for profit, as 'productive', avoiding this complication, without any loss of substance in his argument. Indeed it would have gained in simplicity and therefore in force.[2] His distinctions in treatment between industrial and commercial capital and between industry and distribution as categories of economic activity are, of course, important, but they do not necessitate treating the output of commerce as 'non-productive'. In any case, Marx's categories were chosen to fit his highly simplified model of a capitalist economy, and while he would therefore be justified in excluding, shall we say, personal services, so as to limit his analysis to the employment of labour for profit, there is no reason why these categories should be retained for administrative planning to-day.[3] If it is to be more than a methodological device, the division of work into 'productive' and 'non-productive' must ultimately refer to the social usefulness of the work concerned.

Marx's treatment of 'non-productive' labour was not, of course, original. He took it from the classical school, who based their distinction on the difference between goods and services.[4] We should to-day call work that is 'unproductive' according to Adam Smith 'consumer's services' and 'administrative services'. Marx also included as 'unproductive' part of the production and distribution of goods. Actually, however, the question of what work produces 'value' is hardly relevant. On close inspection it seems that Mr. Minc does not in fact intend to measure 'value' despite his opening definition. The total 'value', being exclusively a function of labour time, would give no aggregate of use to planners. 'Value' would vary only with the 'productive' labour force, the hours it worked, and its composition (inasmuch as skilled labour time incorporates more 'value' than unskilled). It would be unaffected by changes in productivity,[5] and would therefore be quite useless for administrators who plan how to use those increases in productivity which all governments hope will be the means of attaining their economic objectives. What Mr. Minc really proposes to obtain is the total quantity of output, valued in financial terms, but restricting such output to the categories *productive of 'value'*. The distinction is an important one, for if 'value' is not to be measured, there is no obvious reason why even an orthodox Marxist should restrict his measurement to the categories of 'value'.

---

[1] *Capital*, vol. iii, p. 346. [2] Joan Robinson, *Essay on Marxian Economics*, p. 24 n.

[3] When reading orthodox Marxist contributions one continually notices the use of quotations from Marx in place of theoretical argument, and one wishes it would somehow be possible to return and convince Marx of the need for very great care before he reached decisions that would be treated as binding precedents by men in important positions in subsequent generations—despite warnings by such authorities as Engels and Lenin that Marx had not discovered eternally valid truths.

[4] See, for example, Adam Smith, *Wealth of Nations*, Book II, chap. iii.

[5] Cf. Marx (*Capital*, vol. i, chap. xii): 'The value of commodities is in inverse ratio to the productiveness of labour.' This does not conflict with the statement above, for Marx means here the value of each unit of output. If this falls as productivity rises, total value is constant.

What meaning will Mr. Minc's aggregate have? What will one be able to say if Mr. Minc's aggregate is 10 per cent. higher one year than the next, assuming we are working, as Mr. Minc wants to, in fixed prices?

It will not mean that average individual 'productivity' has increased by that amount, if population is assumed constant, because labour may merely have shifted from, say, the armed forces into transport. Nor does it mean that 'welfare', however defined, has increased to that degree, because the production of patent medicines, say, may merely have increased at the expense of medical services. The association between movements in the British aggregates and 'productivity' or 'welfare' is tenuous enough. A strong case could be made against considering movements in the British aggregates to be associated with 'welfare' at all, even in a sense limited to the welfare provided by the 'market sector' of the economy, and even assuming comparability of individual utilities, because the price structure, which is the implicit weighting structure for quantity relatives, is the product of a grossly unequal distribution of income. However, in the short period, if there are no great changes in income distribution, an increase in real income measured on the British and American definitions will have more content in 'welfare' terms in some sense than a similar movement in Mr. Minc's aggregate.[1]

The meaning of Mr. Minc's aggregate to himself is that it is the 'fundamental income' which helps to support the 'non-productive' activities listed above. Mr. Minc quotes Marx as saying: 'Productive workers create a material basis for the maintenance, i.e. for the existence, of non-productive workers', whose services in Mr. Minc's words 'are an act of consumption, of distribution of the national income'. Productive work is the 'foundation', non-productive the 'superstructure'. Mr. Minc quotes a passage from Marx which states that a fall in industrial output implies a fall in the purchase of services—'thus it is clear that the work of a teacher or doctor does not create a direct fund from which it can be paid'. This is a *non sequitur*, since it could also be argued that the fall in a teacher's or doctor's income (e.g. as a result of government retrenchment) would mean smaller purchases of goods, and hence lower output of industry. The position is formally symmetrical. It is, however, true that an increased output of goods was historically the pre-condition of the expansion of the 'non-productive' activities, and that in the long run it is the growth of industry which has permitted the expansion of such services. However, what we are presumably concerned with is short-period planning, and over the short period such 'non-productive' activities are scarcely dependent on the level of industry. Indeed, as the output of industry rises, several 'non-productive' services (book-keeping, administration, education, trade, science, &c.) are the necessary parallel of such increases and also the essential pre-condition of the further advance of industry.[1] This general principle is true for both capitalist and socialist economies, though the nature of services will of course differ considerably—under capitalism taking the form of a multiplication of distribution outlets, personal and commercial services, and under socialism an expansion of administrative services.[2] It could even be argued that certain administrative services were 'fundamental' and the rest of the economy the 'superstructure'; and as far as personal services such as domestic, catering, &c., are concerned, these are in no way less 'fundamental' than luxury goods, considered historically or in the context of current problems of administration.

---

[1] Though *whose* welfare is increased will require further analysis of price and income.

Mr. Minc adds some comments on the American national income estimates, some of which are factually incorrect[3] and some of which are attacks on the Americans for 'double counting'. There is really no content in these attacks, since they follow automatically from his definition of 'income'. If, like Colonel Chinstrap, I define welfare exclusively in alcoholic terms, then all other output is merely a 'cost' of producing alcohol (or an 'intermediate product' in modern terminology), and those who include such auxiliary output in their national income estimates are guilty of grave 'double counting'. It would, however, hardly be proper all the same to wax as ironic over the consequences of differences in definition as Mr. Minc and Mr. Kochanovskov do in discussing, for example, the 'absurd notion' of including defence,[4] or as Mr. Kochanovskov does in attacking Dr. Stadnik for including the wages of clerks in the Czech national income.

---

[1] Mr. Minc himself calls for a great expansion in economic research services as a priority. e.g. 'A large, as yet untouched, problem without which no planning is possible in Poland, is to decide what place is occupied by trade in Poland's mixed economic structure, how to classify trade . . . .'

[2] According to Mr. Minc, Soviet public administration other than that 'whose task it is to manage directly production' is considered 'non-productive'!

[3] e.g. 'The non-Marxist, bourgeois formula is: national income is the sum total of individual incomes, plus the undistributed profits of enterprises', 'the Americans carefully deduct from the sphere of material production, taxes and interests'. Mr. Kochanovskov also thinks that the Americans deduct taxes from profits, and hence distort the picture of income distribution, because 'companies publish their incomes after tax'. 'Jako důsledek toho, že akciové společnosti vykazovaly důchod po zaplacení daně, . . . vzniká v národnim důchodě zmenšený podíl soukromého kapitálu a zvětšený podíl mezd a platů . . .' (*Sociálni Revue*, 1948, No. 9–10, p. 236). Marx would never have been guilty of this sort of inaccuracy.

[4] Mr. Minc also objects, quite consistently from his point of view, to Marshall's adding together (in an example) the incomes of a landowner, the landowner's secretary, and the secretary's servant into a 'sum total' of incomes (since the incomes of the last two are 'secondary'). This viewpoint is, oddly enough, reflected in a passage by Mr. Quintin Hogg,

The exclusion of 'non-productive' services advocated by Mr. Minc should not be confused with the exclusion of 'intermediate' products advocated by Professor Kuznets[1] and recently supported by Mr. Hicks.[2] The differences are several and important. Mr. Minc would exclude from the national income all consumer's services as well as commercial services whether in industry or in trade and all services provided by the government, whereas Professor Kuznets would not. Where they would agree is in excluding defence and certain services provided by the government to industry (e.g. part of the road-building and commercial services, &c.). In addition it has long been a wish of Professor Kuznets to subtract 'outlays representing costs to consumers of membership in our business civilisation',[3] examples of which he gives as 'the standards of an important executive' and elsewhere 'subways, expensive housing, etc.'[4] But he lately seems to have come round to believing such a deduction statistically impracticable, at least till 'a practical analysis of personal consumption expenditures has been made'.

Mr. Minc's definition is very much narrower than Professor Kuznets's or any of the others currently proposed, and corresponds most nearly to the concept of Net *Industrial* Product. This is an aggregate with a certain, though a limited, content, and its use depends on the amount of data available about the rest of the economy. It is, however, misleading for Mr. Minc to apply the word 'national' to what is produced by only part of the nation, and the widespread use of this aggregate would make Eastern European definitions incomparable with others, in itself a very serious disadvantage for economists of all nationalities.

That its use as an aggregate is slight is not itself an overwhelming dis-

---

dealing with Mr. Strachey's use of national income figures to show what income per head would be if it were equally divided. Mr. Hogg assumes a community of one very rich man with £10,000 per annum, and 26 other people living off the rich man and off each other. 'National income' is made up as follows:

	£
Very rich man	10,000
The rest: Receipts from the very rich man . .	10,000
Receipts from one another . . .	10,000
Total . .	

Mr. Strachey would say, according to Mr. Hogg, that this £30,000 means just over £1,000 per head. 'And he is quite surprised when the under-gardener points out to him that all there was to distribute was £10,000 in 2½ per cent. consols.' (*The Case for Conservatism*, p. 179.) Mr. Hogg underestimates the intelligence of under-gardeners—there is the £10,000-income of the rich man, *and* the £20,000 worth of services provided by the rest of the community, to distribute. If the £10,000 of the very rich man were distributed among the others, there would be more services sold between members of the community and consequently higher income and welfare for all the other 26. Mr. Hogg will have to think up a better justification of property income than this quasi-Marxist one.

[1] Vide, e.g., 'National Product in Wartime' (Part I), 'National Income: A Summary of Findings' (Part IV); *Economica*, Feb. 1948, May 1948; *Review of Economics and Statistics*, Aug. 1948.

[2] *Economica*, Aug. 1948, p. 164.

[3] *Review of Economics and Statistics*, vol. xxx, p. 157.

[4] *Studies in Income and Wealth*, vol. i, p. 37.

advantage. The story of national income theory over the past few years has been very much that of liberation from aggregates. The main interest of an aggregate lies in its relation to its components rather than in its intrinsic uses. It is significant that the great development of official National Income statistics came as a result of the war, and quite deliberately, as far as Britain and the United States were concerned, as a guide to the allocation of resources in war-time.[1] The value of the aggregate is that it is in the short period virtually a constraint on the movements of the components (Consumption, Investment, Defence, &c.) and hence indicates what are the limiting possibilities in the allocation of resources between the major alternative objectives.[2]

Now the test of national income definitions lies in their usefulness—as anyone will agree who is concerned to avoid separating theory from practice. The overwhelming objection to Mr. Minc's formulation is that it greatly hinders practical analysis of economic problems, since his restricted aggregate cannot at all easily be divided to show the uses of the country's resources (between total Consumption, total Investment, &c.). This might not be very important in Poland's present economic condition, for reconstruction needs impose overriding priorities on the allocation of resources, but it becomes important as soon as the choice between, say, expansion of consumption and of investment moves into the realm of controversy—even if such controversy is only feasible within the higher ranks of the Polish Communists. A further major objection, particularly from a Marxist viewpoint, is that it is impossible to describe the distribution of national income between social classes, since money incomes generated in large sections of the economy are excluded from 'national income', and it is also moreover impossible to show what proportion of the national product is devoted to capitalist consumption of services.[3] Finally, it is not possible to move easily from national income on Mr. Minc's definition to an estimate of total purchasing power, or to show the relation between total output and purchasing power, which is a grave drawback, especially in a country subject to inflationary pressure.[4]

It might be possible still to defend Mr. Minc's formulation if his aggregate were easy to estimate. Estimators are always free to argue, are indeed bound to argue in some degree, that their estimates conform to what can be measured rather than to what should be measured, their defence resting basically on the tendency for movements in economic activities to be highly correlated. But Mr. Minc does not have this defence. In the first place, a particular difficulty in his definition lies in the need to exclude the 'unproductive' work paid by industry, such as book-keeping and commercial work, and it would involve vastly more statistical work than the definition previously used in Poland. To the question of what constitutes 'value', 'volumes have been devoted in the Soviet Union, based on research into various enterprises and various phenomena of economic activities', according to Mr. Minc, thereby supplying an interesting sidelight on the functions of economic research in the Soviet Union, and he advocates for Polish national income work 'an apparatus of thousands of people who would supply facts and figures'. It is, secondly, unfortunately not necessarily true that the movements of the excluded (service) sectors will be highly correlated with Mr. Minc's 'national income'.

---

[1] Vide, e.g., Milton Gilbert, 'War Expenditures and National Production', *Survey of Current Business*, Mar. 1942; and Keynes, *How to Pay for the War, passim.*

[2] Mr. Kochanovskov quite rightly points out that national income analysis in this way 'eases the work of legislative bodies and reduces the tension of party struggle'. 'Znalost národního důchodu vnáší jasno do celého státního zřízení; ulehčuje práci zákonodárných sborů, snižuje napětí stranického zápasu.' (*Sociální Revue*, 1948, No. 9–10, p. 233.)

[3] If one were to use Mr. Minc's terminology, his methods amount to 'an apology for capitalism'.

[4] Mr. Minc pours scorn on the idea that 'demand' determines production. It is self-evident that in a planned economy the market demand has much less influence than in a *laisser-faire* economy, but surely it is undeniable that monetary demand (as evidenced in the black market, for example) is a factor that any planner must take into account.

Still, the decision of Eastern European governments to adopt Mr. Minc's definition would not necessarily in itself be disastrous for international economic analysis, or for internal planning—provided that sufficient information on the rest of the economy, excluded as 'unproductive', is given to enable national income aggregates on other bases to be devised. There is no one ideal definition suitable for all purposes, as Mr. Minc implies. For 'welfare' comparisons one may well wish to eliminate many of the services Mr. Minc would exclude, while 'productivity' calculations or the descriptions of resource allocation necessitate a more 'gross' concept; and the analysis of effective demand requires a total of personal incomes, including those not derived from the production of any goods or services. The great merit of the official British and American estimates, particularly of the 'national accounts' which have recently been added, is that they generally enable the most appropriate total for any purpose to be selected. The activities of its economists and the technique of planning are the Polish Government's own affair. But the analysis of international trade, the economic summaries produced by international organizations,[1] and the comparison in descriptive economics of the structure or development of different countries, all demand internationally comparable estimates, and it is greatly to be hoped that Eastern European governments will appreciate the advantages for themselves and others of making sufficient information available.

INSTITUTE OF STATISTICS,
OXFORD.

---

[1] e.g. the *Survey of the Economic Situation and Prospects of Europe* (Economic Commission for Europe), or *Survey of Current Inflationary and Deflationary Tendencies* (United Nations Department of Economic Affairs), which both use comparative national income analysis.

A NOTE ON PRICE AND QUALITY

By N. H. LEYLAND

It would be very useful if we could use price differences as adequate indicators of quality differences, for example, in productivity comparisons. This note examines the extent to which this has proved possible in the case of one small industry. I use differences in average prime costs as a measure of quality differences. This narrows down the meaning of the words 'quality differences' to those attributes of a commodity which are the result of more expensive workmanship, better materials, and greater use of variable capital. The definition explicitly eliminates differences in 'quality' which can properly be imputed to locational advantages, the imperfection of the market, salesmanship, and so on.

The Industry Problem

It is very doubtful whether it is possible to speak with any precise meaning of the bicycle industry.[1] The largest and economically most important section produces the everyday bicycle, the smaller section produces light-weight machines. Even this division cannot be maintained. The mass-assembly section is once more producing light-weights, whilst the light-weight section does not produce ordinary machines. There is thus an asymmetrical relation between the mass-assembly section where the large oligopolistic firm[2] is predominant, and the 'hand-made' section where the firm is typically small.[3] It is very unlikely that the markets for ordinary cycles and light-weights overlap in any significant degree and the cross-elasticity of demand is probably extremely small, for the two kinds of machine are required and designed for completely different purposes. The number of firms which would be included in the bicycle industry would vary according to the qualifications for inclusion. If we make the market for cycles in general define the limits, there are two industries, for there are at least two markets. From the point of view of the tube manufacturer there is one industry, including all firms who make bicycles, since steel-tubing is the basic component. On the other hand, the greater part of tube output goes elsewhere. If we define the industry by reference to the process and technique of manufacture we are no better off. For some firms use mass-assembly methods on purchased components, others manufacture their own components and buy only tubing, some light-weight manufacturers assemble by hand without jigs, and so on; finally there is at least one firm which combines all three methods.

---

[1] The Census of Production, probably for quite different reasons, lumps cycle manufacturing with motor-cars.

[2] Leak and Maizels (*J.R.S S.* 1945) give 70 per cent. of employment in the three largest firms. My guess is that these largest firms produce over 80 per cent. of output.

[3] In view of the predominantly small-firm structure of the light-weight section, this section cannot overlap with the mass-assembly section.

The Quality Problem

If for the purpose of this note we use the market division (splitting off from the large firm that small section of their capacity devoted to light-weights) it is quite sensible to speak of the light-weight group or light-weight industry. It is characteristically an assembly industry drawing basic components from a few large firms (e.g. tubes, bearings, rims, malleable iron), who specialize in this field, though for some the greater part of their output goes to other 'industries'. In view of this fact and the relatively simple process of manufacture which must be done by hand, there are good grounds for taking average prime costs as constant for variations in output and as the same for different firms (this was confirmed wherever possible by observation). There is in practice a hierarchy of prime costs according to the grade of machine produced, and depending upon the quality of raw materials and the number and skill of the man-hours incorporated in the finished frame. This does pass back, as it were, some of the problem on to the component manufacturers and it is therefore necessary to examine these prices. In fact it is extremely likely that the quality criteria established above are satisfied. An appeal to the market for components will establish this point. The buyers of components are producers equipped with considerable technical knowledge and the product at the semi-manufactured stage has to pass only technical judgements. Their price must therefore reflect, to a considerable degree, quality differences: for, if there were no genuine quality differences, only the cheapest would be bought.

The market for light-weights does appear at first sight highly imperfect. It seems to contain all those elements of product differentiation which Mrs. Robinson and Mr. Chamberlin have made familiar. This is confirmed to a superficial view by a wide variation in prices, which could only be supported in the long run if each firm had attached to it its own clientèle. On closer examination this no longer appears to be true.

By correcting quoted prices for known differences in prime costs of incorporated components the dispersion of prices is considerably reduced. Two grades of frame were selected for study, the cheapest grade and an 'average' grade, determined by the median product of a firm judged to be typical. (There is a considerable element of personal judgement in this selection, though other methods which were tried yielded similar results.) In the case of the former the coefficient of variation[1] is about 2 per cent., and all the known prime cost differences have not been eliminated. It may be that the cheaper frame is affected by price leadership of the mass-assembly industry, for this is where the sections overlap most. This would therefore appear as a special case if it did not prove possible to eliminate the dispersion of prices of other grades. For then the dispersion would be the result of market imperfection.

---

[1] $\frac{\sigma p}{\bar{p}} \times 100$.

Unfortunately it has not proved possible to eliminate all cost variations from the prices of the average grade. The crude figures yield a coefficient of over 7 per cent. Refinement on the basis of known and established prime cost differences reduce that figure to 4 per cent. The cost differences that are known and allowed for, all reduce the coefficient, since all bring the individual observations nearer to the mean. It is reasonable to assume (and nowhere contradicted by observation) that eliminating the other quality differences would reduce the price variation (those not allowed for were: different labour intensities, different lugs, finishes, delivery costs, and any others which could not be located).

It was necessary to adopt this method of reducing the dispersion within a grade in view of the difficulty involved in collecting labour costs. It represents a first approximation to correcting the dispersion about the average for all machines, which proved impossible because of the ever-increasing number of variables to be corrected. Thus the problem of reconciling the means of grades has not been solved, and the theoretically possible case remains that though market imperfections may not be significant within each grade, they may be significant between grades. I can only say that on the basis of observation I do not think this is the case.

Price Determination

This seems to lead to the conclusion that where there is a considerable degree of interdependence between sellers a standard price will be established for a standard product and deviations from this price will to a great extent be representative of genuine quality differences. In markets which prima facie appear to exhibit all the phenomena of great imperfection the conclusion may well apply that the average price is the standard price and for analytical purposes it may then be legitimate to speak of *the* market price. The conditions under which this result may be expected are two:

1. The buyers must be skilled buyers. In the case I have examined this is certainly true. A large proportion of the purchases of light-weights is made by people who are interested in the machine from a technical point of view. They can judge the merits of different 'qualities' on technical grounds. Purchases are planned in advance, since they usually represent a high percentage of that part of income which they are prepared to spend on 'recreation'. My impression from a large number of conversations with

light-weight enthusiasts is that they get a good idea of the standard or market price and then consider the genuine quality differences they can afford to incorporate in their particular machine. In general one might expect this to be true when buyers require the product of a firm for further manufacturing purposes.

2. Gross margins as a percentage of selling price must not vary widely for each grade of product. This is likely to be the case where the cross-elasticity of demand between grades is high. This condition reinforces the first, and depends upon the existence of the kind of market I have described.

BRASENOSE COLLEGE,
OXFORD.